DOCUMENT RESUME

ED 432 650 UD 033 067

AUTHOR	Turnbull, Brenda; Welsh, Megan; Heid, Camilla; Davis, William; Ratnofsky, Alexander C.
TITLE	The Longitudinal Evaluation of School Change and Performance (LESCP) in Title I Schools. Interim Report to Congress.
INSTITUTION	Policy Studies Associates, Inc., Washington, DC.; Westat, Inc., Rockville, MD.
SPONS AGENCY	Department of Education, Washington, DC. Office of the Under Secretary.
PUB DATE	1999-06-00
NOTE	150p.
CONTRACT	EA96008001
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Academic Achievement; *Compensatory Education; *Disadvantaged Schools; Educational Finance; Educational Practices; Elementary Secondary Education; *Federal Programs; Longitudinal Studies; *Resource Allocation; *School Districts; Urban Schools
IDENTIFIERS	*Elementary Secondary Education Act Title I

ABSTRACT

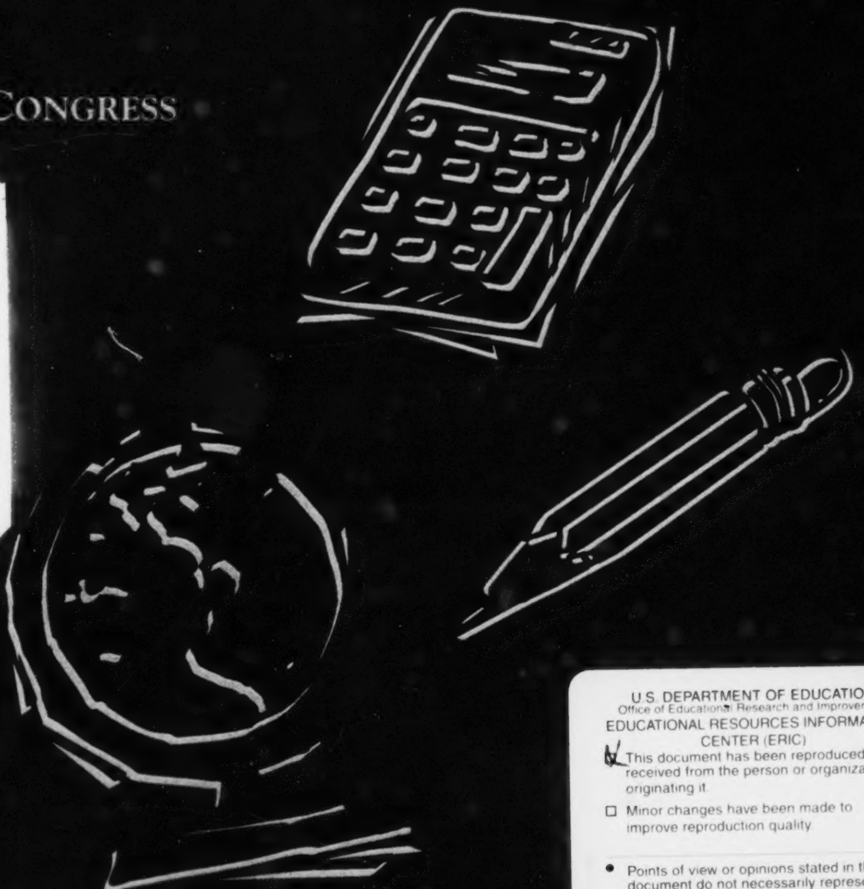
This interim report provides findings on Title I services and policies and on instructional practices and their impact on student performance over the first 2 years of the Longitudinal Evaluation of School Change and Performance (LESCP). The LESCP is being conducted in 71 elementary schools in 7 states and 18 school districts that moved early to implement standards, align assessments, and adopt other elements of reform policy. The study has already identified some teacher variables associated with different rates of growth in student performance in reading and mathematics in high-poverty Title I schools. The study has also identified variation across policy environments in teachers' views and the professional development they have experienced. In the future, the study will focus on other relationships among: (1) interventions intended to influence classroom practice; (2) actual practices reported by teachers; and (3) results for students. Results to date confirm that there are no simple answers, but this research will continue to consider Title I impact on educational reform and academic achievement. Two appendixes contain additional information on study methodology and a discussion of the relationship between student test score gains and teacher responses to survey items. (Contains 83 tables and 10 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 432 650

THE LONGITUDINAL EVALUATION OF SCHOOL CHANGE AND PERFORMANCE (LESCP) IN TITLE I SCHOOLS

INTERIM REPORT TO CONGRESS



Prepared for:
Office of the Under Secretary
U.S. Department of Education

Under:
Contract No. EA96008001

WESTAT
1650 Research Boulevard
Rockville, MD 20850

June, 1999

UD033067

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

THE LONGITUDINAL EVALUATION OF SCHOOL CHANGE AND PERFORMANCE (LESCP) IN TITLE I SCHOOLS

INTERIM REPORT TO CONGRESS

Brenda Turnbull, Ed.D.
Megan Welsh
Policy Studies Associates

Camilla Heid, Ed.D.
William Davis, Ph.D.
Alexander C. Ratnofsky, Ph.D.
Westat

Prepared for:

Office of the Under Secretary
U.S. Department of Education
Contract No. EA 96008001

June, 1999

TABLE OF CONTENTS

Chapter Page

EXECUTIVE SUMMARY xiii

1. INTRODUCTION 1

1.1 The Conceptual Model 2
1.2 Data Sources for This Interim Report and for the Full Study 3
1.3 The Sample: Policy Environments 5
1.4 The Sample: Title I Schools 8

2. STUDENT PERFORMANCE ON LESCP'S TESTS 13

2.1 The Standardized Tests 13
2.2 LESCP Test Scores Available 14
2.3 Cross-Sectional Analyses 16
2.4 The Longitudinal Sample 19
2.5 Relationship Between the Stanford 9 and State Assessments 20
2.6 How We Analyzed the Data on Student Performance 22
2.7 Conclusions 24

3. READING CURRICULUM AND INSTRUCTION 27

3.1 The Reading Curriculum: Skills Emphasized 28
3.2 Frequency and Duration of Instructional Activities 31
3.3 Work with Students of Varying Ability 38
3.4 Teachers' Self-Reported Preparation in Reading/Language Arts 41
3.5 Teachers' Response to Standards-Based Reform in Reading/Language Arts 43
3.6 Families and Schools 46

4. MATHEMATICS CURRICULUM AND INSTRUCTION 51

4.1 Topics and Skills Emphasized 52
4.2 Teachers' Instructional Activities 55
4.3 Students' Instructional Activities 61
4.4 Teachers' Self-Reported Preparation in Mathematics Teaching 68
4.5 Teachers' Response to Standards-Based Reform in Mathematics 68
4.6 Families and Schools 72
4.7 Conclusions 76

TABLE OF CONTENTS (CONTINUED)

Chapter Page

5. POLICY ENVIRONMENT AND PROFESSIONAL DEVELOPMENT: AN INITIAL EXPLORATION OF INFLUENCES ON TEACHERS 77

5.1 Policy Environments in the LESCP Districts 77
5.2 Professional Development and Preparation in Reading 83
5.3 Professional Development and Preparation in Mathematics 88
5.4 Relationship Between Professional Development and Teachers' Reports on Policy Instruments 92
5.5 Relationship Between Professional Development and Curriculum and Instruction 97
5.6 Conclusions 99

APPENDIX A 101

APPENDIX B 107

REFERENCES 129

List of Tables

Table

1. Grades tested, by year of data collection 15

2. LESCP sample sizes 15

3. Missing data in LESCP open-ended tests 16

4. Cross-sectional data on test and subtest performance 17

5. National and urban norms 18

6. Sample size and mean scores for LESCP longitudinal sample 20

7. LESCP teacher response rates for fourth grade in 1998 23

8. Comparison of LESCP students with national norms on 1998 fourth-grade tests or subtests: percentage of the LESCP sample that falls into the bottom quarter nationally 24

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table Page

9. Curriculum emphasis in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain 29

10. Curriculum emphasis in reading/language arts: the percentage of fourth-grade teachers who emphasized a skill 30

11. Frequency and duration of student instructional activities in reading/language arts: percentage of fourth-grade teachers who reported having students engage in instructional activities 32

12. Total exposure to student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain 33

13. Frequency of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain 34

14. Duration of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gains 35

15. Frequency of student instructional activities in reading/language arts: percentage of fourth-grade teachers who reported having students engage in instructional activities by frequency 36

16. Duration of student instructional activities in reading/language arts: percentage of fourth-grade teachers who reported having students engage in instructional activities 39

17. Instructional strategies in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gains 40

18. Instructional strategies in reading/language arts: percentage of fourth-grade teachers who report use of instructional strategies with students of varying ability 41

19. Teacher preparation in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain 42

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table Page

20. Teacher preparation in reading/language arts: percentage of fourth-grade teachers who report their level of preparation to use a variety of instructional strategies 43

21. Policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain 44

22. Policy instruments in reading/language arts: percentage of teachers who are familiar with and implementing standards-based reforms in their classrooms 45

23. Perceived appropriateness of policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gains 46

24. Parent involvement and students ready to learn (reading): relationship between fourth-grade students' gains 47

25. Percentage distribution of responses for fourth-grade teachers and all teachers to parent involvement and students ready to learn survey items (reading) 48

26. Topical coverage in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 53

27. Cognitive demand in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 54

28. Cognitive demand in mathematics: percentage of fourth-grade teachers who emphasized a skill 56

29. Teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain 57

30. Frequency of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 58

31. Duration of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 59

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table Page

32. Frequency of teacher instructional activities in mathematics: percentage of fourth-grade teachers who reported teacher instructional activities 60

33. Duration of teacher instructional activities in mathematics: percentage of fourth-grade teachers who reported instructional activities 61

34. Student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 62

35. Frequency of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 63

36. Duration of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 64

37. Frequency of student instructional activities in mathematics: percentage of fourth-grade teachers who report having students engage in instructional activities 66

38. Duration of student instructional activities in mathematics: percentage of fourth-grade teachers who reported having students engage in instructional activities 67

39. Teacher preparation in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 69

40. Teacher preparation in mathematics: percentage of fourth-grade teachers who report their level of preparation to use a variety of instructional strategies 70

41. Policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 71

42. Policy instruments in mathematics: percentage of teachers who are familiar with and implementing standards-based reforms in their classrooms 73

43. Perceived appropriateness of policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains 74

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table | Page

44. Parent involvement and students ready to learn (mathematics): relationship between fourth-grade teacher's response and fourth-grade students' gains 75

45. Percentage distribution of responses for fourth-grade teachers and all teachers to parent involvement and students ready to learn survey items (mathematics) 76

46. Change in teacher familiarity with and adherence to policy instruments: reading 79

47. Change in teacher familiarity with and adherence to policy instruments: mathematics 80

48. Teachers' familiarity with policy instruments in reading: percentage of teachers familiar with and implementing policy instruments, by district policy environment 81

49. Teachers' familiarity with policy instruments in mathematics: percentage of teachers familiar with and implementing policy instruments, by district policy environment 82

50. Percentage of teachers of reading who responded that they are "very well prepared" in selected teaching strategies 84

51. Percentage of teachers of reading who responded that they participated in "no" professional development on selected topics 85

52. Percentage of teachers of reading who responded that their professional development activities were designed to support reform efforts to a "great extent" 87

53. Percentage of teachers of reading who responded that their knowledge and skills were enhanced to a "great extent" as a result of professional development experiences during the past year 87

54. Percentage of teachers of math who responded that they are "very well prepared" in selected teaching strategies 88

55. Percentage of teachers of math who responded that they participated in "no" professional development on selected topics 90

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table Page

56. Percentage of teachers of math who responded that their professional development activities supported reform efforts to a "great extent" 91

57. Percentage of teachers of math who responded that their knowledge and skills were enhanced to a "great extent" as a result of professional development experiences during the past year 92

58. Correlation between familiarity with policy instruments and the extent to which the professional development was designed to support the policy environment or the extent to which the professional development enhanced the teacher's knowledge and skills 95

59. Correlation between the extent to which the policy instrument is reflected in the curriculum and the extent to which the professional development was designed to support the policy environment or the extent to which the professional development enhanced the teacher's knowledge and skills 96

60. Correlations between mathematical competencies and the extent to which the professional development enhanced the teachers' knowledge and skills ... 99

A-1. Variance of LESCP longitudinal sample score gains by school, teacher, and student 102

A-2. Percentage of variance of LESCP longitudinal sample score gains 103

B-1. Curriculum emphasis in reading/language arts *for low-achieving students*: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 108

B-2. Total exposure to student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 109

B-3. Frequency of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 110

B-4. Duration of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 111

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table | Page

B-5. Instructional strategies in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 112

B-6. Teacher preparation in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 113

B-7. Policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 114

B-8. Perceived appropriateness of policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 115

B-9. Parent involvement and students ready to learn (reading): relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 116

B-10. Topical coverage in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 117

B-11. Cognitive demand in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 118

B-12. Teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 119

B-13. Frequency of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 120

B-14. Duration of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 121

TABLE OF CONTENTS (CONTINUED)

List of Tables (Continued)

Table | Page

B-15. Student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 122

B-16. Frequency of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 123

B-17. Duration of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 124

B-18. Teacher preparation in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 125

B-19. Policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 126

B-20. Perceived appropriateness of policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 127

B-21. Parent involvement and students ready to learn (mathematics): relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students 128

List of Figures

Figure

1. Conceptual framework .... xiv, 3

EXECUTIVE SUMMARY

The 1994 amendments to Title I of the Elementary and Secondary Education Act introduced major changes in that program's policies. Increasingly, Title I is designed to work in concert with state and local reforms, especially those reforms based on aligned frameworks of standards, student assessment, curriculum, and professional development. Over this period of the law's authorization, states are expected to move in the direction of standards-based reform, and districts and schools are expected to use their Title I funds to enable the students served to meet challenging state standards.

This is the interim report of a large study, the Longitudinal Evaluation of School Change and Performance (LESCP), designed to measure changes over time in selected students, classrooms, and schools participating in Title I. This summary describes the study's purposes and design, then presents highlights of its findings to date.

Study Purposes

The study's research questions focus on curriculum and instruction, student performance, and the effects of the 1994 Title I amendments. They are as follows:

1. To what extent are changes occurring in what is being taught in reading and mathematics in grades K-5 in the classrooms in the study?
2. To what extent are changes occurring in how instruction is being delivered?
3. To what extent are students showing changes in performance?
4. How do recent revisions in Title I contribute to these changes?

The study places these questions within a conceptual framework that shows a logical chain of connections from Title I policy to student performance (Figure 1). With research questions largely focused on curriculum, instruction, and student performance, LESCP focuses its attention most closely on the right-hand side of this framework. However, it was also designed to explore the policy environment set by states and districts and the implementation of instructional programs by schools, so that it can test the contribution of these possible influences on changes in curriculum, instruction, and performance. In this regard, the study pays especially close attention to the extent to which policy and program implementation conform to Title I policy: for example, it has documented state and local policies on standards and aligned curriculum, assessment, and accountability, and it has gathered detailed data on teachers' participation in professional development.

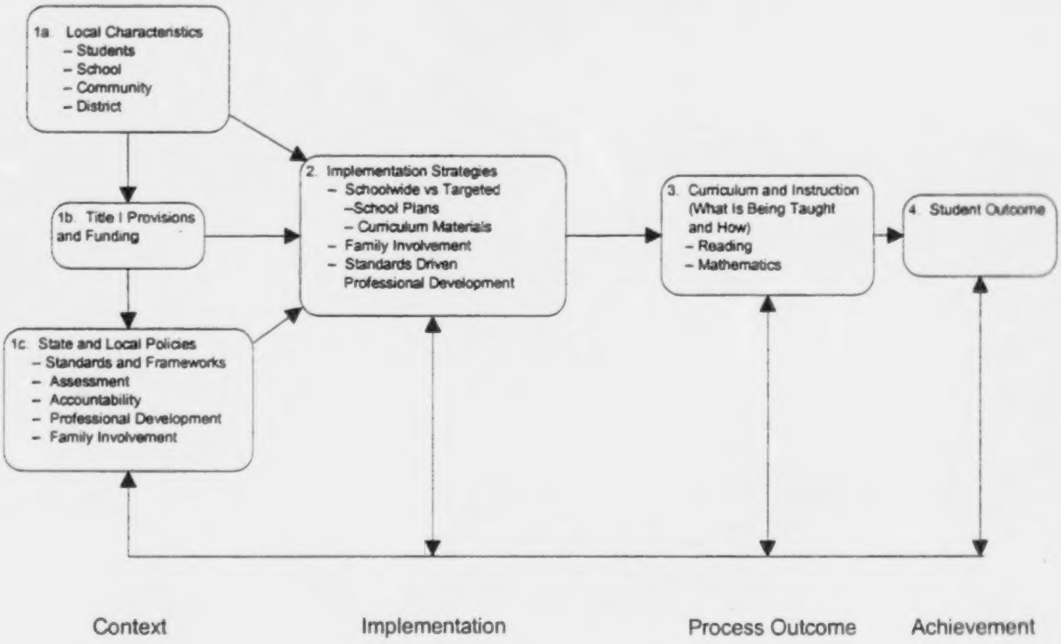

1a. Local Characteristics
– Students
– School
– Community
– District
1b. Title I Provisions and Funding
1c. State and Local Policies
– Standards and Frameworks
– Assessment
– Accountability
– Professional Development
– Family Involvement
2. Implementation Strategies
– Schoolwide vs Targeted
–School Plans
– Curriculum Materials
– Family Involvement
– Standards Driven Professional Development
3. Curriculum and Instruction (What Is Being Taught and How)
– Reading
– Mathematics
4. Student Outcome
Context
Implementation
Process Outcome
Achievement

Figure 1. Conceptual framework

Study Sample and Methods

The LESCP study has gathered data in the spring of 2 school years in 71 high-poverty elementary schools around the country, all receiving funds under Title I. A third year of data collection is taking place in spring 1999. Although the schools are not statistically representative of the program or the nation as a whole, they do provide a substantial amount of information about ways in which implementation of Title I is currently unfolding, especially at the classroom level, and on the relationships between classroom practice and achievement.

The LESCP schools are nested in 18 districts within 7 states—policy environments with a disproportionately high level of activity in standards-based reform. Although there was variation in the extent to which the states had enacted standards-based policies in 1996, when the sample was drawn,

BEST COPY AVAILABLE

that variation has diminished as states have moved forward on standards, assessment, and accountability for schools. Thus, the differences in state policies across the sample now have less to do with the presence or absence of a standards-based approach and more to do with (1) the length of time systems have been in place and (2) the sequence of policy enactment (e.g., some have had high-stakes assessment in place for some time, while others began with standards and professional development). Similarly, the 18 districts do not vary dramatically in whether they have reform activity. As a group, they tend to reflect their states' relatively active policies, and in fact some have moved more rapidly and aggressively than their states to establish accountability systems.

Several sections of this report identify teacher and school differences across "high-reform" and "low-reform" policy environments. This analysis focuses on the extremes within this sample of districts. At the "high" end of the continuum are the four districts with the most detailed policies on standards-based reform (most of these policies originating at the state level); at the "low" end, are four districts that have moved as slowly as their states have allowed them to in embracing a standards-based approach. Thus, taking into account both state and local policy, we have identified districts in which schools experience especially high levels of formal policy on standards, assessment, and accountability, and districts in which there are fewer policies on these subjects (in comparison with the rest of the sample). Standards, assessment, and accountability are not the only possible dimensions of reform, but they are the ones on which our analyses have focused so far.

This study's investigation of services in schools focuses at the classroom level. Most of the schools in the sample—58 of 71—operate schoolwide Title I programs, where "Title I services" are not necessarily a discrete set of experiences offered to participating children but where participation in Title I brings an expectation that the school will do what it takes to enable all children to meet challenging standards. This study's data show that principals of most of the schools say that Title I plays a major role in providing extra help to low-achieving students, in supporting partnerships between parents and schools, and in professional development; more than half also use Title I programs to support extended time for students. About three-quarters of the schools are funding the salaries of teachers as part of their Title I programs; just under two-thirds have paraprofessionals.

Data collected for this study include the following, featured in this interim report:

- **Standardized tests** administered to third- and fourth-grade students in spring 1997 and to fourth graders in spring 1998; this design permits the measurement of achievement growth by individual students over a year. (In 1999, fourth and fifth graders are being tested, yielding data on an additional year of growth for individual participants and for schools.)

- **Surveys of teachers** in every K-5 classroom in every school, with extensive questions about curriculum and instruction in reading and mathematics, as well as questions about professional development, their own level of preparation, and school-home partnerships.
- **Interviews with principals and district Title I coordinators, and collection of school and district documents** to support the analysis of policy environments.

Additional qualitative and quantitative data can be analyzed in future reports across the full 3 years of the study:

- **Data on student performance on state or local assessments.** These data will permit us to analyze the relative performance of schools within states, although they are not comparable across states, and the publicly released data do not shed light on variation by classrooms or growth of individual students over time. Performance on these tests can also be compared with the achievement data that we collected by administering a national standardized test.
- **Observations of reading and mathematics lessons in selected classrooms,** illustrating the differences in types of instruction.
- **Qualitative data gathered from focus groups with teachers and parents.** These data will be especially useful in identifying the organizational dynamics of improvement in those schools that are registering good gains in student performance.
- **Qualitative data on policies and programs of family involvement** gathered at the state, district, and school levels, including focus-group data from parents.

Student Performance on LESCP's Tests

This report analyzes the association between classroom curriculum and instruction, as reported by teachers, and gains in performance by individual students who took the study's standardized tests in both years. The Stanford 9 achievement tests assess both reading and mathematics with both open-ended (constructed-response) and closed-ended (multiple-choice) tests. The closed-ended tests have separate subtests of vocabulary and comprehension, in reading, and problem solving and procedures, in mathematics.

As a group, the students participating in this study performed somewhat below national and urban norms in both years. The students tested as fourth graders in 1998 performed about 0.8 grade levels higher than those tested as third graders in the previous year, showing somewhat less progress than would be expected over a full year. The proportion meeting higher proficiency levels ("superior" or "solid academic performance") held fairly steady across the 2 years. The subset of students who were tested in both years outperformed their more transient peers but still provide us with a large database with variation in achievement and growth.

For purposes of this study's analyses, the Stanford 9 data offer several advantages and one noteworthy drawback. The advantages include (1) having data on individual student growth across years as well as school performance trends, (2) being able to link student performance with teacher reports, and (3) having data that are comparable across states. It is true, however, that Title I charges schools with improving student performance on state assessments that are aligned with challenging state standards. To the extent that a state's standards might be quite different in content from the knowledge and skills measured by the Stanford 9 tests, these standardized tests are not an ideal proxy for the performance that Title I seeks to boost. And, indeed, where we could directly compare the performance of two successive fourth grades in the same school on both a state test and a Stanford 9 test, the direction of the change between years was the same in only about two-thirds of the cases. For this study's future analyses of trends at the school level, the data available from state assessments will be important.

Our analysis delved deeply into the variation in student growth associated with reported teacher behaviors, looking separately at different subgroups within classrooms. For each aspect of curriculum and instruction that teachers described in their surveys, we tested whether there was a systematic relationship with the rate of student performance growth. The data collection and analysis paid particular attention to the possibility of differences across students in the class depending on their prior achievement—a subject of intense policy interest in Title I over the years. Thus, teachers were asked numerous specific questions about the curriculum and instruction that they provided to students who began the year with skill levels lower than their classmates'. In the analysis of student growth, we looked separately at those students whose third-grade scores placed them in the bottom quarter of their incoming fourth-grade class with respect to the particular reading or mathematics test or subtest. Growth registered by these "bottom-quarter students," as well as by their "top three-quarters" classmates, was analyzed in relation to teacher responses about curriculum and instruction.

Classroom Practices Associated with Student Growth in Fourth-Grade Reading

There were some relationships between fourth-grade teachers' survey responses and their students' growth in reading, especially with regard to the frequency of instructional activities, the teachers' own level of preparation in several instructional techniques, and the way families prepared their students for learning. However, the kinds of practices associated with better student growth in reading were not necessarily the dominant ones in this sample of classrooms, and there was little systematic change in the direction of these practices over the 2 years of data collection.

- **The particular skills that a teacher emphasized in the curriculum (e.g., comprehension, vocabulary, or phonics) showed few positive relationships with student growth.** For example, comprehension was an area of emphasis for most teachers, and more so with "typical" students than with low achievers—yet its only apparent benefit was for students in the bottom quarter. Emphasis on phonics increased over the 2 years of the study, but the teacher's report of emphasis on phonics was not significantly associated with student growth. (We should underscore again that these results were obtained with fourth graders; the curriculum that is effective with early learners may not work as well for these older students.)

- **However, more positive relationships emerged for students' exposure to relatively demanding instructional activities—the kinds of things they did in the classroom—and especially for the frequency of these activities.** For example, bottom-quarter students tended to gain more with those teachers who had them read materials of at least a paragraph, read materials in the content areas, talk in small groups about what they had read, and work at a computer. More positive relationships to growth were found when teachers used these instructional activities *often* than when they used them *for a long duration in each lesson.*

- **Over the 2 years of the study, the change in teachers' practice was generally *not* in the direction of practices associated with better student growth.** For example, the frequency of practicing word attack increased, although this activity was not associated with growth; the frequency and duration of use of reading materials in the content areas decreased, although this activity *was* associated with growth.

- **Teachers appeared to be good judges of their own skills.** All students made better gains in the classrooms of teachers who believed they were well prepared to teach heterogeneous groups and use a variety of assessment methods; bottom-quarter students' gains were positively associated with teachers who felt well prepared with respect to *any* of the instructional skills asked about. Moreover, teachers' self-reported levels of preparation in these skills increased across the 2 years.

- **Teachers' answers to questions about the policy instruments of standards-based reform showed no particular relationships with students' growth.** Although there were increases in teachers' reports that they were familiar with standards, frameworks, and assessments and were following these policy instruments, such reports were not associated with student growth—perhaps reflecting, in part, the fact that the Stanford 9 differs from state tests.

- **For the bottom quarter of students, growth in performance had a positive relationship with teachers' reports on both parent involvement and students coming to school ready to learn.**

Classroom Practices Associated with Student Growth in Fourth-Grade Mathematics

In mathematics, more relationships were found between teachers' survey responses and their students' growth in performance. Several kinds of teacher responses were associated with better rates of growth in fourth-grade mathematics.

- **Positive relationships were found between the number of lessons taught in each of many mathematical topics and student gains.** This was especially true for students in the top three-quarters of their class.
- **A curriculum that focused on the skills of understanding concepts, solving equations, and solving problems was associated with better gains in problem solving for students in the bottom quarter of their class.** Ironically, teachers were less likely to emphasize understanding concepts and solving problems with these students; however, their emphasis on more demanding cognitive skills with these students did increase over the 2 years.
- **Student growth was positively associated with total exposure to activities calling for active student participation (taking a test, using manipulatives, discussing multiple approaches to solving a problem) rather than more teacher-focused activities (a lecture or presentation at the blackboard).**
- **As in reading, frequent repetition of an activity (using it daily or weekly) appeared to be a good strategy, but remaining with a particular activity for a long time within each lesson could be negatively associated with growth.** Encouragingly, the changes reported between 1997 and 1998 tended to be in the direction of more frequency and shorter duration.
- **Again echoing a finding from reading, bottom-quarter students had better gains with teachers who gave higher assessments of their own preparation to teach mathematics.** Skills that seemed especially valuable were those of teaching heterogeneous groups and taking students' existing skills into account. Teachers tended to report increases in their level of preparation in these skills across the 2 years.
- **More than in reading, there were some positive associations between student growth and teachers' reports that they were familiar with and using standards, assessments, and curriculum frameworks, especially for bottom-quarter students.**

- **There were some positive relationships between bottom-quarter students' gains and teachers' reports about parental involvement or students' readiness to learn.** The relationships were less clear in mathematics than in reading, however.

Policy Environment and Professional Development as Influences on Classrooms

This report also addresses some of the conditions that could be expected to influence teachers' opinions and behaviors, especially the policy environment and the professional development in which teachers participated. For these analyses, we broadened our focus beyond the fourth grade to encompass all the K-5 teachers in the participating schools.

Having arrayed the 18 districts from high to low along a rough continuum of standards-based reform (as described above), we found that the responses of teachers to questions about standards-based reform differed in the expected ways across district policy environments. Teachers in "high-reform" districts were significantly more likely to report knowing and following various policy instruments, such as standards, assessments, and frameworks. The changes over time were interesting as well: the greatest amount of change in teacher responses was found in those districts that did not start out at either the high or the low extreme of the policy environments.

Professional development varied a great deal across districts, although only some of the variation was associated with the gross distinction between high- and low-reform environments.

- One example of a difference was the greater emphasis on learning about assessment in high-reform districts.
- Another was the greater focus on state and district reforms in high-reform districts, whereas the school's own reform plan was more often the focus of professional development in low-reform environments.

The overall amount of professional development diminished across the 2 years of the study, at least with respect to the selected topics about which teachers were asked. At least one-fourth of the teachers said they had participated in *no* professional development during the 1997-98 year on each of several topics: content in reading, instructional strategies for teaching reading, content in mathematics, and instructional strategies for teaching mathematics. Still higher proportions had had no professional development on strategies for teaching low-achieving students or strengthening parental involvement.

When asked whether professional development had helped them in each of a variety of ways—such as adapting to standards or assessments, or gaining confidence in using new approaches—fewer than one-fourth of the teachers said it had helped "to a great extent."

Participation in professional development was only modestly associated with *differences* in classroom practices, and it was not discernibly associated with *changes* in practice for individual teachers across the 2 years of the study. Examples of the relationships found—none of them very large, in statistical terms—included the following:

- Teachers who said their professional development had focused on policy instruments such as state assessments also tended to report relatively high adherence to those policy instruments.
- Teachers who reported that professional development had helped them encourage students to collaborate in math class, or had helped them use a range of instructional techniques in mathematics, tended to place greater emphasis on problem solving in their mathematics curriculum. This emphasis was, in turn, related to achievement gains among lower performing students.

Looking Ahead

So far, then, this study has identified some teacher variables associated with different rates of growth in student performance in reading and mathematics in high-poverty Title I schools. These variables will be pursued in future analyses, adding data from fifth-grade teachers and students to the mix when those data become available. The study has also identified variation across policy environments in teachers' views and the professional development they have experienced and has begun to explore other relationships among (1) the interventions intended to influence classroom practices, most notably professional development; (2) the actual practices reported by teachers; and (3) the results for students. Although the results to date confirm that there are no simple, resounding answers, the study will continue to pursue its in-depth, longitudinal assessment of factors associated with greater and lesser success in a realm that has crucial importance to this nation's future: using Title I resources to bolster the fundamental academic skills of children who are growing up in poverty.

1. INTRODUCTION: STUDY PURPOSES, DESIGN, AND SAMPLE CHARACTERISTICS

This Interim Report of the Longitudinal Evaluation of School Change and Performance (LESCP) analyzes variation and changes over time in students' performance and teachers' curriculum and instruction in a set of high-poverty elementary schools. The study's design and analyses are organized around the policies embodied in Title I of the Elementary and Secondary Education Act, as amended in 1994. Ever since its original enactment in 1965, Title I has been intended to improve the learning of children in high-poverty schools, with a particular focus on those children whose prior achievement has been low. Therefore, this study measures changes in student performance in a sample of Title I schools, and its analyses take a special look at those students with initially low achievement. The study also surveys teachers about the classroom curriculum and instruction offered to students, again with attention to any differences in the program offered to those students who begin with comparatively low achievement.

The study's research questions are the following:

1. To what extent are changes occurring in what is being taught in reading and mathematics in grades K-5 in the classrooms in the study?
2. To what extent are changes occurring in how instruction is being delivered?
3. To what extent are students showing changes in performance?
4. How do recent revisions in Title I contribute to these changes?

The final question refers to the new provisions of Title I, enacted in 1994, that strongly encourage states, school districts, and schools to pursue a standards-based approach to educational improvement. The standards-based approach relies on aligned frameworks of standards, curriculum, student assessment, and teacher professional development to set clear goals for student performance and to help organize school resources around those goals. It is an approach that several states and some large districts began to put in place earlier in the 1990s. Several large Federal programs, prominently including Title I, adopted the philosophy of standards-based reform in 1994.

This chapter describes the conceptual model that organizes the study's data collection and analysis. It then highlights the particular data sources that have been most thoroughly explored at this interim point in the study and describes the additional data that can be incorporated into a future final report. Two final sections of the chapter describe the variation found in the study sample with regard to important dimensions of Title I policy and programs: first, the variation in state and local policies with respect to standards-based reform and second, the Title I program designs found in the study's schools.

1.1 The Conceptual Model

The study's conceptual framework, depicted in Figure 1, directs the analysis toward relationships among context, implementation, instructional process outcomes, and student achievement outcomes. We show the process outcomes at the teacher level as box 3 in Figure 1. These outcomes, constrained to reading and mathematics in this study, are the curriculum (what is being taught) and the instruction (how the material is being taught). Curriculum and instruction are two of the major outcome measures of the study. For all the K-5 classrooms in schools participating in this study, teachers responded to questionnaires about their curriculum and instruction in reading and mathematics in both study years (and are doing so again in the third year).

At the same time, curriculum and instruction are inputs to the student achievement outcomes, represented in box 4 of the figure. Students took the Stanford 9 achievement test in grades 3 and 4 in 1997 and in grade 4 in 1998; this permits us to track the performance gains of individual students over time and also the changes in performance across successive cohorts of fourth graders. Continuing the pattern, students were tested in grades 4 and 5 in 1999 so that an individual student's growth can be followed for an additional year and so that we have tested three successive fourth grades in each school. Student records are linked to teacher records so that we can identify relationships between teachers' survey responses and student performance. A substantial part of this report explores those relationships with the 1997 and 1998 data.

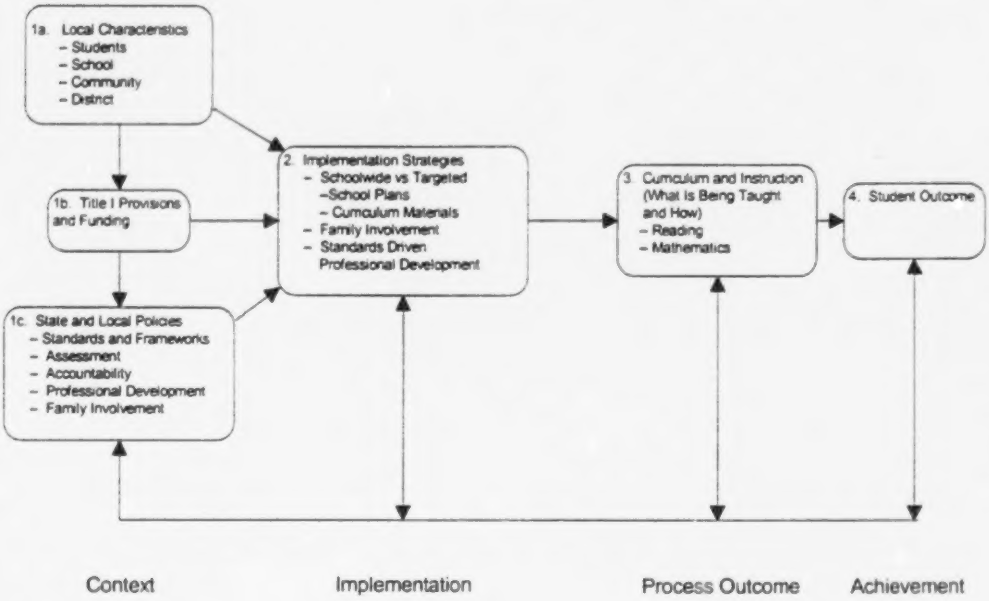

1a. Local Characteristics
– Students
– School
– Community
– District
1b. Title I Provisions and Funding
1c. State and Local Policies
– Standards and Frameworks
– Assessment
– Accountability
– Professional Development
– Family Involvement
2. Implementation Strategies
– Schoolwide vs Targeted
–School Plans
– Curriculum Materials
– Family Involvement
– Standards Driven Professional Development
3. Curriculum and Instruction (What is Being Taught and How)
– Reading
– Mathematics
4. Student Outcome
Context
Implementation
Process Outcome
Achievement

Figure 1. Conceptual framework

One way to look at the framework and the study is to emphasize the arrows pointing from left to right—to investigate the influence of policies, mediated by implementation, on classrooms and in turn on student performance. This chain of influence is important to pursue, and we have begun to do so in the analyses reported here. At the same time, we recognize that the LESCP is a longitudinal study of a dynamically changing educational system. We therefore have shown a feedback loop pointing from right to left in Figure 1. As administrators, teachers, and evaluators observe outcomes, we would expect this information to influence state and local educational policies, implementation strategies, and curriculum and instruction. Our analyses must be mindful of the likelihood that some of the policies and practices found at the district or school level are the *result* of prior teacher and student performance at the same time as they are an *influence* on emerging classroom and individual outcomes.

1.2 Data Sources for This Interim Report and for the Full Study

The report is based on two rounds of data gathered in spring 1997 and spring 1998 from students and teachers. With a third and final round of data collection occurring in spring 1999, the

BEST COPY AVAILABLE

LESCP study is collecting repeated measures of students' performance, teachers' reported behavior and opinions, and the school's policy environment in 71 schools. These schools, all of which receive funds under Title I, are nested in a purposively selected sample of 18 districts in 7 states. The schools are not statistically representative of high-poverty schools in the nation as a whole, in their states, or even in their districts. However, the study provides a rich database that permits the analysis of differences across students, classrooms, schools, and policy environments at any one time and also across school years. The first 2 years' data have provided an opportunity to begin exploring trends and relationships in students' performance linked with their teachers' survey responses.

This report focuses on the right-hand side of the conceptual model, beginning to draw relationships between Box 4, student growth as measured on the Stanford 9 tests and subtests, and Box 3, classroom curriculum and instruction as reported by teachers. These analyses draw most heavily on two of the study's data sources:

- The tests administered to students who were in grade 3 in 1997 and grade 4 in 1998
- Surveys completed by fourth-grade teachers in 1998 regarding topics that include their classroom curriculum and instruction in reading and mathematics, their knowledge and preparation with regard to standards-based reform and instruction, and their professional development over the past 12 months

To place these preliminary findings in perspective, the report also describes trends in teachers' survey responses over time, for fourth-grade teachers and also for all K-5 teachers.

This report also takes an early look at selected policy and program variables that might be expected to affect classroom curriculum and instruction. If we found that they did, it would then make sense to investigate their effects on students' progress. Within Box 2 of the conceptual model, "Implementation," this report takes a particular focus on professional development. Professional development is considered an important communication channel between policy and the classroom: teachers can only act on what they know, and it is plausible that they are most likely to know about—and perhaps act on—standards-based reform when they have participated in professional development driven by standards. The study has collected data from individual teachers regarding their own participation in professional development over 12 months, providing a good basis for investigating relationships with a large sample of cases.

Finally, this report uses data collected from the school districts about the policies enacted by either the district or the state that reflect standards-based reform. These policies are depicted in Box 1c of the conceptual model. Again, because time has not yet permitted a comprehensive look at all relationships among all variables, the study team has chosen to focus on the dimensions of reform most closely tied to student performance and classroom curriculum and instruction: standards and curriculum frameworks, assessment, and accountability.

Additional data sources are part of this study but have not yet been fully incorporated into its analyses at this interim point. They include the following:

- Data on student performance on state or local assessments (a different data source for Box 4, "Student Outcomes"). These data will permit us to analyze the relative performance of schools within states. The advantage is that these are the assessments most salient to students and teachers. The disadvantages are that they are not comparable across states and that they do not give us a means of looking at either (1) variation by classrooms or (2) growth of individual students over time.
- Observations of reading and mathematics lessons in selected classrooms as a supplemental source of insight into Box 3, "Curriculum and Instruction." Although a single observation does not give a reliable basis for assessing curriculum and instruction over a full academic year, the classroom observations will illustrate the differences in types of instruction.
- Qualitative data on schools, gathered from interviews with principals, focus groups with teachers and parents, and school plans. The qualitative data will help inform the analysis of Box 2, "Implementation." These data will be especially useful in identifying the organizational dynamics of improvement in those schools that are registering good gains in student performance—what has created a sense of urgency about improvement, and what has helped the school improve?
- Qualitative data on policies and programs of family involvement at the state, district, and school levels, along with focus-group data from parents (most of whom were selected for their high levels of participation in their children's schooling). These data can be used to develop profiles of vigorous efforts to involve parents as educational partners and to investigate the relationship between such efforts and student behavior (as reported by teachers) and performance.

1.3 The Sample: Policy Environments

This study looks in depth at a purposive sample of state and local policy environments rather than using a larger, nationally representative sample of Title I schools. To assess school and classroom responses to standards-based reform, the study focuses on states and districts that enacted standards-based reform some years ago.[1] The states vary in their approach to reform—for example, some put high-stakes assessment in place years ago, while others began their reforms with a process of developing content standards. Of the seven states in the sample, five were arguably embarked on some version of standards-based reform in 1996 when the sample was drawn. Although the other two states were doing less with standards-based reform in 1996, they have moved in that direction over the course of the study, and one of them has moved quite rapidly. This leaves the study with less variation at the state level than was originally expected. The 18 participating districts present a similar pattern: none is untouched by standards-based reform, although they vary in the alacrity and thoroughness with which they have enacted each of several kinds of standards-based policies, both in response to state requirements and on their own initiative. In short, the LESCP schools are subject to some variation in the kinds of policies enacted by their states and districts—but it is important to recognize that all have been subject to some policy activity in standards, assessment, or accountability.

So that we could look at the variation in teachers' survey responses by policy environment, we used documents provided by the districts' offices to create ratings for each of the 18 districts on several indicators of standards-based reform policies in 1998. We focused this analysis at the district level in order to capture both state and district policy as embodied in district documents. Taking this approach was necessary because our sample deliberately included pairs of states and districts that had initially taken different policy stances on standards-based reform. The sample included, for example, districts that state officials described as reluctant to implement an aggressive standards-based agenda initiated at the state level. It also included districts that had independently established their own standards-based framework in the absence of such a framework statewide.

The following indicators were used to distinguish between those districts with policies highly reflective of standards-based reform and those at the other end of the continuum:

Standards and Aligned Curriculum Materials:

- District has content or performance standards in at least reading and math, either state or locally determined

---

[1] The study could not meet its mandate simply by looking at schools' responses to the provisions of the 1994 law because the timeline for the study does not mesh well with the timeline for the law's implementation. For example, Title I does not require full implementation of standards-based accountability for schools until the school year 2000-2001, 2 years after LESCP data collection ends. Thus, only in those states and districts that had already enacted standards-based reform some years ago—before the Title I provisions were enacted—would it be possible to expect widespread, classroom-level effects as early as the LESCP data collection period.

- District standards are clearly linked to state and/or national professional standards
- District and school improvement plans are standards-based
- Curriculum guides are aligned with state and/or local standards

Assessment:

- Some assessments are performance-based at each developmental level
- District reports assessment data in terms of its own or the state's proficiency levels (e.g., novice, proficient, satisfactory, etc.)
- District and school improvement plans set goals linked to performance standard proficiency levels

Accountability:

- District has defined adequate yearly progress for Title I according to the state's standards or it has built on the state's definition to derive its own
- District has policies to reward or sanction schools on the basis of their achievement of the district and/or state standards
- District periodically reports school and district achievement status to the public in readable and understandable formats using data disaggregated to two or more of the Title I categories
- In addition to student achievement data, other accountability indicators report on students' cognitive and noncognitive progress
- District and school improvement plans reflect the district's and state's accountability expectations

Raters assigned a score from 2 to 0 (full, partial, or zero) to each district with respect to each indicator. With four indicators for standards and curriculum, three for assessment, and five for accountability, the possible scores therefore ranged from 24 to 0. The range actually found among all 18 districts was from 22 to 9. For this analysis, we looked at teacher responses according to where their districts were in this range (at the high or low end, or somewhere in the middle).

The four "high-reform" districts had scores of at least 20, and the four "low-reform" ones had scores of 12 or below. The distribution of these eight outlier districts across states suggests the importance but also the limits of state policy. Of the four high-reform districts, three were in one state with a long-standing and comprehensive reform approach; the fourth was an urban district in a state that has used assessment and accountability as key elements of its reforms. However, in no state was every participating LESCP district uniformly rated as "high-reform." The four low-reform districts were scattered across four states. As a group, they reflected some combination of permissive state policy (where local control or gradual implementation over several years has been an important principle) and limited local capacity to generate policies and documents for their schools. No state in this sample had districts uniformly rated as "low-reform" because districts could and did enact their own frameworks of standards, assessment, and accountability.

In analyzing the data gathered from teachers and schools, we have looked for variation and trends by policy environment. We would expect to see more attention to standards-based reform among teachers, for example, in those districts where more standards-based policies are in place. Later sections of this report explore those relationships.

1.4 The Sample: Title I Schools

The 71 schools in the LESCP sample all receive Title I funds, and most have very high levels of poverty. We describe here the overall patterns and variation found in the way the schools implemented their Title I programs.

Of the 71 schools, 58 were operating schoolwide programs in 1997-98 (up from 54 in the previous year). This reflects, in part, the high poverty levels of participating schools: 25 schools had more than 75 percent of their students living in poverty; 16 schools had between 50 and 75 percent; and 22 had fewer than 50 percent (but in all cases more than 35 percent).

When asked about the role played by Title I in their school, almost all principals (68 of 71) said it had a "major" role in "helping provide extra instruction for low-achieving students." This response was chosen by 56 of the 58 schoolwide-program principals and 12 of the 13 principals in targeted assistance schools. More details are available about the specific instruction funded in the 13 targeted schools, where it had to be separately accounted for: reading instruction was a focus in 12 schools, mathematics in 11; grades 2-4 were served in all 13 schools, grade 5 in 12, grade 1 in 11, and kindergarten in 8. Title I teachers provided instruction in 11 schools and Title I aides in 11. Pullout designs were found in 12 schools, small in-class groups in 11, and in-class team teaching in 9.

Title I was said to play a major role in "involving parents as partners in their children's education" in a smaller but still substantial number of schools: 56 of 71 overall, including 48 of the 58 schoolwide programs and 8 of the 13 targeted programs. A difference across school types was in the use of parent-involvement coordinators. The use of Title I funds for that position was reported in 12 schoolwides but none of the targeted programs.

"Expanding professional development opportunities for classroom teachers" was a major role for Title I in 43 schools, including 41 of the 58 schoolwide programs. This was a relatively unusual role in the targeted-assistance schools, reported by just 2 of 13 principals—perhaps because it sounded like using Title I for general aid. When asked an open-ended question about the elements of the school's Title I program, nearly identical proportions of schoolwide and targeted-assistance principals said professional development was one use of Title I funds. Presumably the professional development was directed to Title I staff in the targeted-assistance schools.

Extended time was funded by Title I in 40 of the 71 schools. The designs included summer programs in 31 schools, before- or after-school care or instruction in 26 schools, and an extended year in 11 schools (with several schools reporting more than one of these choices).

An open-ended question about the use of Title I funds provided some additional data. When principals were asked what Title I supported in their schools, about three-quarters of them mentioned teachers; this included all but one of the principals of targeted-assistance schools. Sixty-two percent mentioned paraprofessionals, with similar percents in schoolwide and targeted programs. Professional staff other than teachers—counselors, social workers, nurses—were reported in seven schools, all of them schoolwide. Materials were mentioned in 62 percent of schools and computer technology in 29 percent, with no large differences between schoolwide and targeted schools.

Thirteen schools were implementing a model of comprehensive school reform. This total included five implementing Accelerated Schools, four implementing Success for All, and four implementing the Comer School Development Program. All of these, except one Accelerated School, operate schoolwide programs.

Using the analysis of policy environments described above, we looked for systematic differences in Title I designs and other aspects of the school program across high- and low-reform environments. Although the small numbers of schools suggest caution in interpreting these data, some differences did emerge between the 15 schools in high-reform environments and the 9 in low-reform environments:

- Schools in high-reform environments were more likely to use their Title I funds to pay for academic staff, including teachers in 80 percent of these schools and paraprofessionals in a partially overlapping 80 percent. By contrast, just 44 percent of the schools in low-reform environments mentioned teachers as part of the program, and 33 percent mentioned paraprofessionals.

- Although schools in high-reform environments used professional development, technology, or an extended school day or year, they did not use Title I funds to pay for them.

- When asked how much they used "innovative technologies" as a strategy for reform, 5 of the 15 principals in high-reform environments said "to a great extent," outnumbering the 1 of 9 principals in low-reform environments who gave this answer.

- An extended day was used as a reform strategy "to a great extent" in 7 of the 15 schools in high-reform environments and an extended year in 5. An extended day was used "to a great extent" in one of the nine schools in low-reform environments and an extended year in none. (Several more principals in low-reform environments reported using their Title I funds for extended time but did not rate this as a reform strategy that their school was implementing to a great extent.)

- No school in a high-reform environment was implementing a model of comprehensive school reform, although 13 schools overall were doing so.

- There were some differences in principals' reports about the impetus for changing curriculum and instruction, depending on their policy environments. Curriculum frameworks were a "great" impetus in 9 of the 15 schools in high-reform environments but 1 of 9 in low-reform environments. Textbooks, on the other hand, were less likely to be a "great" impetus in high-reform environments (3 of 15 schools versus 4 of 9).

Similarly, we looked for systematic differences between the 13 schools that had been identified for improvement under Title I and the 54 that had not (excluding from this analysis the 4 in which principals said they did not know whether they had been identified).

- Those identified for school improvement were less likely to say they were paying for teachers in their Title I programs—just 42 percent of the identified schools, compared with 81 percent of nonidentified schools.

- The identified schools were more likely to use Title I funds for professional development (three-fourths versus one-half of the nonidentified schools).

- They were also more likely to employ other professionals (counselors, social workers, etc.), who were found in one-third of the identified schools but just 6 percent of nonidentified schools.

- When asked about the impetus for changing curriculum and instruction, the factor that was mentioned at a higher rate in schools identified for improvement than in other schools was "changes in student demographics," which was an impetus for change in three-fourths of the identified schools but about one-fourth of those that were not identified.

This descriptive analysis of the study schools reveals some of the variation in Title I programs, which can inform future analyses of performance trends at the school level. At this point in the study, however, we have focused much more extensively on understanding the performance of individual participating students over time. This takes advantage of the unique strengths of this particular study's design and methods: unlike most other current studies in the policy arena, it takes a close look at students and their classroom environments. We turn next to a description of the study's data on student performance.

2. STUDENT PERFORMANCE ON LESCP'S TESTS

A major source of information in this study is the student testing conducted with third and fourth graders in the spring of 1997 and with fourth graders in the spring of 1998 in the 71 LESCP schools. This chapter describes: the standardized tests, the students who took each of the tests in the spring of 1997 and 1998 (as well as the extent and causes of missing data), overall results for all the students and for the subset of students who were tested in both years, and a comparison of performance trends at the school level between the standardized tests and states' own assessments. This is a description of Box 4, Student Outcomes, in our conceptual framework. The chapter then describes the analytic procedure by which we investigated the relationship between student gains and classroom curriculum and instruction (or, how we studied the arrow running from Box 3 to Box 4).

2.1 The Standardized Tests

Test scores were obtained using the Stanford 9 achievement tests. Separate scores were obtained for each of eight tests and subtests (four tests, plus two subtests within each of two of those tests) in the spring of 1997 and 1998:

- Overall closed-ended reading
  - Closed-ended vocabulary
  - Closed-ended comprehension
- Open-ended reading
- Overall closed-ended mathematics
  - Closed-ended mathematics problem solving
  - Closed-ended mathematics procedures
- Open-ended mathematics

The overall closed-ended reading score is a composite of the vocabulary and comprehension scores, while the closed-ended mathematics score is a composite of the problem-solving and procedures scores. The Stanford 9 is a norm-referenced achievement test. According to the publisher, the mathematics subtests align with the National Council of Teachers of Mathematics (NCTM) standards, and the reading comprehension subtest aligns with the National Assessment of Educational Progress (NAEP).

The multiple-choice reading test is composed of two subtests, vocabulary and comprehension, at the grades administered in LESCP. The vocabulary subtest assesses vocabulary knowledge and skills with synonyms, context clues, and multiple word meanings. The reading comprehension subtest uses a reading selection followed by multiple-choice questions to measure modes of comprehension (initial understanding, interpretation, critical analysis, and process strategies) within the framework of recreational, textual, and functional reading. The open-ended reading test contains a narrative reading selection in the recreational reading content cluster followed by nine open-ended questions that measure initial understanding, interpretation, and critical analysis.

The multiple-choice mathematics test is composed of problem-solving and procedures subtests. Five processes are assessed in the problem-solving subtest: problem-solving, reasoning, communication, connections, and thinking skills. Concepts of whole numbers, number sense and numeration, geometry and spatial sense, measurement, statistics and probability, fraction and decimal concepts, patterns and relationships, estimation, and problem-solving strategies are measured. The procedures subtest covers number facts, computation using symbolic notation, computation in context, rounding, and thinking skills. The open-ended mathematics assessment presents nine questions or tasks around a single theme. Ability to communicate and reason mathematically and to apply problem-solving strategies are assessed. The content clusters for the open-ended mathematics test are number concepts, patterns and relationships, and concepts of space and shape.

The number of students for whom we have test scores varies by the test because not every district had its students take each component test. Both the math and reading open-ended tests included all districts in the LESCP study. However, one district did not participate in the closed-ended math test, while two districts did not participate in the closed-ended reading test.

2.2 LESCP Test Scores Available

In this report, we analyze LESCP test scores from data collected during the spring of 1997 and 1998. Table 1 shows the basic sources of data studied here. For 1997, we have scores for the third and fourth grades, while for 1998 we have scores for the fourth grade. We pay particular attention in the analysis to the cohort of students who were third graders in the spring of 1997. Because we have repeated measurements on many of these students, we can measure score growth with a reliable baseline score.

Table 1. Grades tested, by year of data collection

Year		
Spring 1997	Spring 1998	**Spring 1999**
3rd		
4th	**4th**	**4th**
		5th

Much of our analysis focuses on the cohort of students who were third graders in the spring of 1997. In addition to the population of all test takers, we identified a subset of students for further analysis. These were the students who were tested in both years. We call this the longitudinal sample. In contrast to the population of all students in the cohort, this group may be more stable. We know at least that they were tested in the same school in the spring of the third and fourth grades.

Table 2 shows the total number of third- and fourth-grade LESCP students tested for each of the eight tests and subtests in the spring of 1997 and 1998. The minimum number of students for any test, grade, and year is 2567. This is an appreciable sample and should allow us to make reliable conclusions.[2]

Table 2. LESCP sample sizes

Test or Subtest	**3rd Grade**	**4th Grade**	
	1997	**1997**	**1998**
Reading Closed Ended	2813	2692	2567
Vocabulary	2827	2712	2893
Comprehension	3225	3158	3060
Reading Open Ended	3646	3535	3438
Math Closed Ended	3226	3073	2987
Problem Solving	3285	3150	3050
Procedures	3254	3105	3006
Math Open Ended	3723	3503	3400

---

[2] National percentile and mean estimates are based on samples of size 4000 to 5000.

Table 3 documents the causes of missing test scores for two of the eight tests and subtests: open-ended reading and math (the results vary slightly by test, but these two tests are representative). It shows that approximately half of the students who were on the roster in 1997 ended up taking *both* tests.

Student test scores were missing for a variety of reasons, such as students' changing schools between third and fourth grade, lack of parental approval to take the test, absence on the day of the test, etc.; the breakdown of these causes of missing data is shown in Table 3. For example, it shows that a total of 373 students were ineligible for the third-grade reading test in 1997. The primary causes of ineligibility to take the test were disability, 154, and limited-English proficiency, 118. (The schools were instructed to exclude only those students whom they would also exclude from taking their regular assessment for Title I purposes.) Next, the table shows that 751 eligible third graders did not take the test. The reasons were the following: parental refusal, 387, absent, 159, and miscellaneous causes, 166. Of the 3646 students who completed the third-grade test, 841 (or 23 percent) transferred before the spring 1998 test. Of the 494 students who missed the 1998 test, the primary causes were parental refusal, 231, absent, 110, and miscellaneous causes, 64.

Table 3. Missing data in LESCP open-ended tests

	Open-ended Reading		**Open-ended Math**	
	Number	**Total**	**Number**	**Total**
On roster, 3rd grade 1997		4770		4770
Ineligible 3rd grade test 1997	373		328	
Eligible 3rd grader who missed 1997 test	751		719	
Took 3rd grade test in spring 1997		3646		3723
Not on 4th grade roster (transferred)	841		850	
On 4th grade roster but missed 1998 test	494		550	
Longitudinal sample size		2311		2323

2.3 Cross-Sectional Analyses

This section compares the performance of LESCP students with national and urban reference groups and with proficiency levels identified by the test publisher. On average, students in the LESCP sample of schools scored below national norms and urban norms in both years and grades tested.

Table 4 shows the cross-sectional data on test and subtest performance for the entire LESCP sample for 1997 and 1998. The data are shown in several forms: overall mean scores, the national percentile and grade-equivalent that these means represent, and the percentage of LESCP test takers who performed at particular "competency levels" on each test or subtest in each year. These levels are described as corresponding to the kinds of performance levels that Title I encourages for state assessment data (e.g., "excellent," "proficient," and the like).

Table 4. Cross-sectional data on test and subtest performance

	Third Grade 1997						
Test or Subtest	**Mean Score**	**Natl Percentile of Mean Score**	**Grade Equiv. of Mean Score**	**% Level 1: Below Satisfactory**	**% Level 2: Partial Mastery**	**% Level 3: Solid Performance**	**% Level 4: Superior Performance**
Reading Closed Ended	602.3	38	3.4	32%	40%	24%	4%
Vocabulary	597.7	41	3.5	22%	29%	31%	18%
Comprehension	603.2	38	3.3	39%	39%	17%	5%
Reading Open Ended	574.9	37	3.2	37%	32%	18%	12%
Math Closed Ended	592.0	43	3.5	30%	46%	21%	4%
Problem Solving	602.0	43	3.5	30%	36%	28%	6%
Procedures	578.1	44	3.6	29%	42%	21%	9%
Math Open Ended	582.4	32	2.9	29%	36%	25%	9%
		Nation=	3.8				
	Fourth Grade 1997						
Test or Subtest	**Mean Score**	**Natl Percentile of Mean Score**	**Grade Equiv. of Mean Score**	**% Level 1: Below Satisfactory**	**% Level 2: Partial Mastery**	**% Level 3: Solid Performance**	**% Level 4: Superior Performance**
Reading Closed Ended	623.4	35	4.1	35%	39%	19%	7%
Vocabulary	624.8	38	4.3	24%	33%	28%	15%
Comprehension	620.1	35	3.9	41%	32%	19%	9%
Reading Open Ended	598.0	39	4.2	36%	41%	17%	5%
Math Closed Ended	614.0	39	4.4	33%	39%	22%	6%
Problem Solving	616.7	43	4.6	26%	40%	28%	7%
Procedures	610.1	38	4.4	47%	27%	19%	7%
Math Open Ended	590.1	23	3.5	43%	38%	14%	6%
		Nation=	4.8				

Table 4. Cross-sectional data on test and subtest performance (continued)

				Fourth Grade 1998			
Test or Subtest	**Mean Score**	**Natl Percentile of Mean Score**	**Grade Equiv. of Mean Score**	**% Level 1: Below Satisfactory**	**% Level 2: Partial Mastery**	**% Level 3: Solid Performance**	**% Level 4: Superior Performance**
Reading Closed Ended	621.2	34	4.0	34%	38%	21%	8%
Vocabulary	620.0	35	4.2	24%	32%	28%	16%
Comprehension	619.7	35	3.8	41%	30%	21%	7%
Reading Open Ended	601.9	42	4.3	31%	44%	20%	5%
Math Closed Ended	614.3	39	4.4	34%	39%	21%	6%
Problem Solving	618.2	45	4.4	27%	39%	26%	7%
Procedures	609.1	37	4.4	48%	27%	19%	7%
Math Open Ended	589.9	23	3.5	47%	34%	13%	6%
		Nation=	4.8				

For comparison, Table 5 shows the national and urban norms by test. National and urban norms were taken from the Stanford Achievement Test Series (1996) based on a representative sample of student scores in spring 1995. Because urban means were not available for the open-ended test, urban medians were used. On all tests and subtests, the LESCP students fall below the national norms from 4 to 22 points.[3] Although the LESCP students typically scored below the urban norm, they did score slightly above the urban norm for math procedures in the third grade.

Table 5. National and urban norms

Test or Subtest	**3rd Grade Means**		**4th Grade Means**	
	National	**Urban**	**National**	**Urban**
Reading Closed Ended	613.9	606.6	637.2	634.3
Vocabulary	607.9	601.9	638.0	636.4
Comprehension	617.9	609.8	637.7	609.9
Reading Open Ended	586.4	578.9*	606.0	634.3*
Math Closed Ended	599.5	593.0	624.2	623.5
Problem Solving	608.3	601.4	624.2	623.3
Procedures	588.1	581.5	625.8	625.3
Math Open Ended	602.3	590.4*	612.4	609.1*

* Indicates median rather than mean

---

[3] The conclusions based on the differences from national and urban norms are valid if the national and urban scores have not changed substantially since the norming year, 1995.

Looking across the 2 years of testing for LESCP fourth-grade classes, there were statistically significant gains at the .01 level on Open-ended Reading and statistically significant losses in the subtest of Vocabulary. On other tests and subtests, the performance of the two successive fourth-grade groups was similar.

Although changes in mean scores over a population are important to detect, it is also important to know whether other aspects of the score distribution are changing. For example, there could have been regression to the mean with fewer students scoring very high and fewer students scoring very low. However, our investigation of the score distributions showed that this was not the case. On six of the eight tests or subtests, the distributions remained essentially indistinguishable from year to year. On the other two, where there were changes in mean performance, the gain or loss was distributed across students at all levels of performance.

We note that these cross-sectional results were obtained on two different fourth-grade classes in the LESCP schools. In the next section, we analyze the change within the cohort of students who were third graders in those schools in the spring of 1997. In this group, we can more accurately determine whether any statistically significant changes were due to the educational experiences of participating students during the fourth grade.

2.4 The Longitudinal Sample

For the longitudinal sample, we have a reliable baseline score so we can accurately assess the score gain made between the spring of 1997 and the spring of 1998. In contrast to the use of all test takers, the use of the longitudinal sample for analysis has the following two disadvantages: it reduces the sample size, and it limits the generality of the conclusions to those students who spent both third and fourth grade at the same school. However, the advantage of using the longitudinal sample is the ability of longitudinal studies to distinguish changes in achievement over time for individual students between the third and fourth grades from differences in achievement between two successive fourth-grade classes (Diggle, Liang, & Zeger, 1994). Using the longitudinal sample, we can look at relative learning gains for students who are exposed to varying curriculum content and instructional activities.

LESCP students were tested on grade. If a student progressed from grade 3 to grade 4, we required that the student take the third-grade form of the Stanford 9 test in 1997 and the fourth-grade form in 1998, to be included in the longitudinal sample.

Table 6 shows the number of students in the longitudinal sample and the mean scores for this group for the eight tests and subtests taken in spring 1997 and 1998 and the difference in mean scores between the LESCP longitudinal sample and all the LESCP students in the cohort. Not surprisingly, the longitudinal students scored higher than the other test takers in all eight tests and subtests in both years. The difference in means ranged from 1.7 to 7.7 points.

Table 6. Sample size and mean scores for LESCP longitudinal sample

Test or Subtest	Sample size	Mean 1997	Difference 1997	Mean 1998	Difference 1998
Reading Closed Ended	1648	607.4	5.1	626.9	5.8
Vocabulary	1659	603.6	5.9	627.7	7.7
Comprehension	2030	606.0	2.8	624.3	4.6
Reading Open Ended	2311	579.6	4.7	605.2	3.3
Math Closed Ended	1986	595.1	3.1	618.3	4.0
Problem Solving	2060	604.4	2.4	621.8	3.6
Procedures	2017	582.7	4.6	613.8	4.7
Math Open Ended	2323	584.6	2.3	591.6	1.7

The LESCP longitudinal sample students made approximately the same amount of progress as the nation as a whole during the fourth grade, with two exceptions, both in reading:

- Reading: For the closed-ended test, the LESCP longitudinal sample score gain was significantly less than the national score gain. However, for the open-ended test, the LESCP score gain was significantly more than the national score gain.
- Math: Although the LESCP longitudinal sample score gains were below the national gains, the differences were not statistically significant.

2.5 Relationship Between the Stanford 9 and State Assessments

The analyses conducted for this report, focusing on the student outcomes associated with particular aspects of classroom curriculum and instruction, emphasize trends in individual student performance across years. Student performance on the Stanford 9 gives us performance data in a common metric across all the study's classrooms, specifically associated with students' individual demographic characteristics and their own teachers' survey responses. We do not have a comparable level of detail regarding performance on state tests, and those tests vary across states. However, Title I charges states and school districts with improving performance in relation to state standards, as measured by state assessments. Therefore, it is worth checking how well the Stanford 9 results match up with the results of state assessments.

Comparisons were possible for 34 schools in 3 states that (1) provide school-level performance data for both 1997 and 1998 and (2) assess student performance in the fourth grade on reading and/or mathematics. We made a simple comparison. We combined fourth-grade student performance on both Stanford 9 reading tests into a single measure, then found the difference between fourth graders' performance in that school for 1997 and 1998; we did the same for mathematics. We compared the direction of change (up or down) with the direction of change in that school's fourth-grade performance on its state test in the same subject over the same 2 years. In other words, we compared the direction of change registered by the same school over the same 2-year period on the basis of the performance of the same two fourth grades.

In this way, we compared the direction of change on a total of 47 measures, 17 in reading and 30 in mathematics. (More comparisons could be made in mathematics because one of the states assesses its fourth graders in mathematics only.) The direction of change agreed in 30 of these 47 comparisons, with similar rates of agreement in reading and mathematics (10 out of 17 were the same in reading; 20 of 30 in mathematics). Where the two tests disagreed on the direction of change, the school was usually—but not always—moving up on the state test. Across the board, the following performance trends were found:

- 18 schools moved down on both tests in a particular subject (8 in reading, 10 in mathematics)
- 12 moved up on both tests in a subject (2 in reading, 10 in mathematics)
- 12 moved up on only their state test in a subject (4 in reading, 8 in mathematics)
- 5 moved up on only the Stanford 9 in a subject (3 in reading, 2 in mathematics)

The differences in trends suggest that there are some real differences in the skills measured by the Stanford 9 and the state tests. Where we can choose which test to use—for example, in future analyses of changes in performance at the level of the entire school—we should concentrate on the state tests, because they are the ones for which schools are accountable in the Title I policy system. Ideally, we would have been able to conduct all of the study's analyses with those tests. However, that would prevent us from analyzing student growth because only one state tests students in the same subject every year, and none releases the results for individual students by name. It would also preclude analysis of the effects of within-school variation in classroom curriculum and instruction because results of these state assessments are not publicly reported for classrooms, and it would seriously complicate, if not preclude, analysis across states.

2.6 How We Analyzed the Data on Student Performance

Most of the analyses presented in this report investigate factors associated with the gains made by students in the longitudinal sample. In particular, we focus on the relationship between these students' gains and the classroom curriculum and instruction that their teachers offered them in fourth grade. With a large number of students and classrooms available for investigation, we have been able to identify some aspects of curriculum, instruction, and teacher preparation that are positively associated with the rate of student gain—as well as others that show a negative association with student gain. Here, we summarize the analytic procedure used. Further details are provided in Appendix A.

Each of the students was associated with a fourth-grade primary reading and a primary mathematics teacher. For each teacher in the study (except nonrespondents), we obtained data via a self-administered questionnaire. From the questionnaire, we obtained information on the time spent in the average week on mathematics and on reading, class size, the teachers' familiarity and application of several components of standards-based reform, assessment practices, instructional methods, and curriculum for "typical" and for "lower achieving" students. The questionnaire also included the teachers' perceptions of the involvement of parents in their children's learning, professional development opportunities, and the teacher's background and experience.

Questionnaires were collected for all teachers of grades K through 5 in the 71 schools in the study. In much of the analysis carried out below, we restrict attention to the longitudinal student sample and estimate the impact of the fourth-grade teacher's practices on the student's score gain. If the fourth-grade teacher did not complete the questionnaire, we eliminated the data for all students in the class rather than impute teacher responses. Table 7 shows the LESCP fourth-grade teacher response rate for the reading and math open-ended tests and the number of longitudinal sample students included in these classes.[4] For example, for open-ended reading, Table 7 shows that 86 percent of the fourth-grade teachers completed a questionnaire, and 88 percent of the longitudinal students were in classes taught by these teachers. This table shows that 11 to 12 percent of the longitudinal student's scores needed to be eliminated during analysis of the relationship between teacher practices and test score gains due to failure of the teacher to complete the questionnaire on these two tests, which were representative of the other six tests and subtests.

Table 7. LESCP teacher response rates for fourth grade in 1998

	Open-ended Reading			**Open-ended Math**		
	Total	**Respondents**		**Total**	**Respondents**	
		Number	Percent		Number	Percent
4th Grade Teachers	229	198	86	199	170	85
Longitudinal Students	2311	2036	88	2323	2077	89

Table 7 can be used to compute the average number of longitudinal students per teacher. For responding teachers, the average number of students per teacher is 10.3 (=2036/198) for reading and 12.2 (=2077/170) for mathematics. These numbers are smaller than the overall class size due to requirements on the longitudinal sample (among other things the student must have taken the test in the previous year).

For each of the eight tests and subtests, we split the longitudinal students into two groups: bottom quarter and top three-quarters based on pretest results. The split was carried out for each fourth-grade class separately using the students' pretest results (third-grade scores for each test). We split the students in this way because several questions on the teacher questionnaire were directed at practices with lower achieving students, who were specified as the bottom quarter of the class. We assessed the impact of teaching technique for the low- and high-achieving groups independently by carrying out the same analysis on each. Thus, the student split allows us to determine whether the impact of the particular teaching technique varies with the student's prior achievement level.

A different analysis, included as Appendix B, looks at the relationship between teacher behavior and student gains for those students who are in the "bottom quarter" of *all fourth graders nationally*. Among the students participating in the LESCP study, a disproportionate number—between 28 percent and 55 percent—scored in the bottom quartile on these tests nationally (Table 8). Appendix B presents the results from analyses paralleling those presented in chapters 3 and 4, but with students divided according to their performance in relation to national norms.

---

[4] Only fourth-grade teachers who had at least one longitudinal student were included.

Table 8. Comparison of LESCP students with national norms on 1998 fourth-grade tests or subtests: percentage of the LESCP sample that falls into the bottom quarter nationally

Test or Subtest	Bottom Quarter Nationally
Reading Closed Ended	40
Vocabulary	43
Comprehension	40
Reading Open Ended	28
Math Closed Ended	37
Problem Solving	30
Procedures	39
Math Open Ended	55

Our analytic procedure was one of hypothesis testing: we hypothesized that teachers' response to a particular survey question had no relationship to their students' gains. The analysis used four hierarchical levels, with repeated test scores nested within student, within class, and within school. We tested whether the variation in teacher response to each survey question had a significant relationship with student performance at the end of fourth grade, in an equation that also took into account variation by student (i.e., the student's performance in third grade), class, and school. Where the hypothesis could be rejected, we report a relationship.

2.7 Conclusions

The Stanford 9 tests offer information about several aspects of student performance in reading and mathematics, using constructed responses (on the open-ended tests) as well as multiple-choice items. The LESCP sample as a whole performed below national and urban norms on these tests, and the students' proficiency levels held steady across the 2 years of testing. Those students who took tests in both years had somewhat higher levels of performance than their more transient peers. It is these students, the "longitudinal" group, who are the basis for this report's analysis of the contribution of fourth-grade curriculum and instruction to growth in student achievement. Although these standardized tests measure somewhat different skills from any particular state test—as shown by the fact that in about one-third of cases a school's aggregate performance moved in one direction on the state test of a particular subject and the opposite direction on the standardized test of that subject—they do offer a comparable basis for measuring performance and growth across all the study's classrooms. The analysis of curriculum and instruction in relation to growth is the subject to which we turn next.

3. READING CURRICULUM AND INSTRUCTION

Our analysis of teachers' responses to the survey questions about reading and language arts has focused on identifying the association between student growth and particular teacher behaviors, as reported on our surveys. With the achievement data from those students who were tested as third graders in spring 1997 and fourth graders in spring 1998, we were able to measure student growth on each test and subtest. We could then find relationships between the amount of growth registered by each student and the survey responses gathered from that student's fourth-grade teacher. In the study's conceptual framework, we are looking here at the arrow running between Box 3 (Curriculum and Instruction) and Box 4 (Student Outcomes).

This chapter presents the results of that analysis. Necessarily, at this point in the study, the focus is largely on the fourth grade because that is the only grade for which we currently have both a pretest and posttest score for each student on each test. Therefore, most of the data reported in this chapter are drawn from fourth-grade teachers' responses to survey items administered in 1997 and 1998. In addition to examining the relationship between performance gains and teacher responses, we also studied changes in responses over the 2 years to see whether teachers are increasing the use of instructional strategies associated with better student gains. Later sections of this report describe the variation and trends in reading curriculum and instruction across all K-5 teachers in the 71 LESCP schools. These analyses of trends correspond to Box 3 in the study's conceptual framework.

For each of the questions posed to teachers, we tested the hypothesis that the response to that question has no relationship to student growth, and summarize the results in the tables that follow, for reading in this chapter and for mathematics in the following chapter. The table format deserves a few words of explanation here. *If the hypothesis was not rejected* at the .05 significance level, the *cells are blank* in the tables.[5] If the *test was significant* at the .05 significance level, *plus or minus signs* are shown in the tables. *A plus sign* indicates that increasing the quantity of a particular teacher practice is *related to a significant increase* in test scores. Significant *negative relationships* are indicated by *minus signs*.

---

[5] With the large number of tests carried out here, some significant results will be purely due to chance. However, the overall rate of significant effects for reading was 12 percent, which is significantly above the rate expected due to chance. Unless a teaching activity was statistically significant for a majority of the reading or math tests, we were skeptical of its effectiveness.

Title I has historically emphasized bolstering the achievement of students who are at the greatest risk of academic failure, and this study follows that tradition by looking closely at the curriculum and instruction offered to relatively low-achieving students within each classroom. Most of the analyses presented here were performed separately for each of two groups of students in each teacher's classroom—those whose *third-grade* scores on the test or subtest fell into the bottom quarter *relative to the scores of their own fourth-grade classmates on that test or subtest* and those whose scores fell into the top three-quarters for their fourth-grade class. This analytic division of the "bottom quarter" and "top three-quarters" in each class gave us separate information about students who corresponded to two groups that were singled out in many of the survey questions for teachers, "your lowest achieving students (roughly the bottom quarter)" and "your typical students." This procedure allowed us to examine how teachers differ in their approach to dealing with students of varying skill levels and how curriculum and instructional techniques are working for different groups of students.

This chapter presents the results of the analysis of fourth-grade growth in reading in relation to the following:

- The skills that teachers emphasize in their curriculum
- The instructional activities in classrooms, attending to both frequency and duration of the activities
- Teachers' opinions of their own level of preparation to teach reading, their reported familiarity with standards-based reform, and their assessment of families' contributions in preparing their students to learn

This chapter also shows how fourth-grade teachers responded to the survey questions. It identifies which responses were most frequent, whether teachers gave significantly different answers when asked about their lowest achieving students as opposed to their typical students, and how their responses changed over the 2 years of data collection.

3.1 The Reading Curriculum: Skills Emphasized

This study tried to assess reading curriculum by asking teachers about the skills they emphasized in teaching reading. For whatever reasons, the skills that teachers reported emphasizing with their students did not have strong relationships with student growth (Table 9). The exception was "content area reading strategies," which were associated with low rates of growth on the closed-ended test. Content area reading strategies were not emphasized as much as other skills with either low-achieving or typical students (Table 10).

Table 9. Curriculum emphasis in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain

Teacher Report	**Reading Test or Subtest—Bottom-quarter Students**				**Reading Test or Subtest—Top Three-quarters Students**			
	Open-ended Reading	**Closed-ended Reading**			**Open-ended Reading**	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**		**Overall**	**Comprehension**	**Vocabulary**
Extent of Emphasis, in Teaching Low-Achieving Students on:					**Extent of Emphasis, in Teaching Typical Students on:**			
Comprehension	+							
Vocabulary								
Oral Reading					+			
Content Area Reading Strategies		-	-				-	
Phonics/Word Attack								

Table reads: There was a significant (p<.05) positive relationship between (1) a fourth-grade teacher's self-reported emphasis on comprehension in teaching low-achieving students and (2) the gains made by that teacher's bottom-quarter students on the open-ended reading test.

Table 10. Curriculum emphasis in reading/language arts: the percentage of fourth-grade teachers who emphasized a skill

Subject:	**Degree of Emphasis in 1998 (N=246)**				**Degree of Emphasis in 1997 (N=219)**			
	A Lot	Moderately	Occasionally	No Emphasis	A Lot	Moderately	Occasionally	No Emphasis
How much do you emphasize with all students?								
Comprehension	94	5	1	0*	90*	9*	0	0
Vocabulary	68	27	4	0	67	30	3	0
Oral Reading	61+	27+	12	0*	46	42*	12	0*
Content Area Reading Strategies	58	35	6	1*	53	39	7	0
Phonics/Word Attack	45+*	31*	22*	1	33*	39	25*	2
How much do you emphasize with low-achieving students?								
Comprehension	89	9+	0	2+	84	14	1	0
Vocabulary	66	29	5	0	72	24	4	0
Oral Reading	66+	24	10	1+	55	32	12	1
Content Area Reading Strategies	58	33	6	3+	52	38	9	0
Phonics/Word Attack	63+	22+	14	2	53	34	10	2

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

* Indicates a statistically significant (at the $p<.05$ level) difference between teachers' emphasis with low-achieving students and all students in the same year.

For both "typical" and "low-achieving" students, teachers most frequently reported that they emphasized "a lot": comprehension (94 percent and 90 percent, respectively), vocabulary (68 percent and 67 percent, respectively), and oral reading (61 percent and 46 percent, respectively). Even though the same top three skills were mentioned for both groups, there was one significant difference between the two groups: 63 percent of teachers reported emphasizing phonics "a lot" with their low-achieving students, while 45 percent reported this level of emphasis on phonics with their typical students (Table 10).

There were increases between 1997 and 1998 in the degree of emphasis that teachers placed on oral reading and phonics for both groups of students (with larger increases in responses for typical students) and on comprehension for low-achieving students. Although the increase in focus on reading comprehension aligns with increases in test scores for bottom-quarter students, the increased emphasis on phonics skills does not seem to have benefited students—at least, we detected no relationship between teachers' emphasis on this skill and improved student achievement.

3.2 Frequency and Duration of Instructional Activities

We examined the relationship between total exposure to an activity and gains in reading achievement. Total exposure is a derived variable that we calculated, taking into account both the frequency with which an activity was provided and its duration. Because there appears to be little relationship between the frequency and duration of an activity—that is, the activities that were most frequent did not tend to have either especially long or especially short duration (Table 11)—total exposure gives us a method of examining the relationship between the overall intensity of an activity and student learning.

Generally speaking, better student growth tended to be associated with classrooms in which students' total exposure to critical thinking strategies was high, and their exposure to drill activities, designed for students' rote memorization of facts, was low (Table 12). This is reflected by the positive relationship between gains in test scores and activities such as reading materials of at least one paragraph (for bottom-quarter students), reading content area materials (for both bottom-quarter and top three-quarter students), and students talking in small groups about what they have read (for both bottom-quarter and top three-quarter students). Additionally, a negative association was found between achievement growth and total exposure to practicing phonics or practicing word attack strategies for those students in the top three-quarters of their class.

Table 11. Frequency and duration of student instructional activities in reading/language arts: percentage of fourth-grade teachers who reported having students engage in instructional activities

Instructional Activity:	1998		1997	
	All Students	Low-achieving Students	All Students	Low-achieving Students
Frequency of "Almost Every Day"	**(N=242)**	**(N=243)**	**(N=218)**	**(N=217)**
Read Materials of at Least One Paragraph	90*	81	91*	76
Read Aloud	79+	76+	68	62
Read Books They Choose Themselves	58	58+	53	47
Practice Word Attack	52*	62	44*	54
Read Content Area Materials	46+	43	56*	44
Practice Phonics	40+*	52+	29*	43
Complete Workbooks/ Skill Sheets	36	32	35	40
Talk in Small Groups About What They Have Read	35	38+	27	28
Write About What They Have Read	32	34	29	30
Work at a Computer	29	30	28	28
Duration of "More Than Half a Lesson"	**(N=208)**	**(N=185)**	**(N=195)**	**(N=168)**
Read Materials of at Least One Paragraph	18*	31	26*	39
Read Aloud	34+	41+	50	56
Read Books They Choose Themselves	52	52	55	59
Practice Word Attack	57	56	65	61
Read Content Area Materials	25+*	41	38*	51
Practice Phonics	59+	57	70	62
Complete Workbooks/ Skill Sheets	50	50	53	55
Talk in Small Groups About What They Have Read	53	54	52	57
Write About What They Have Read	35	38	36	42
Work at a Computer	39	45	40	39

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

* Indicates a statistically significant (at the $p<.05$ level) difference between teachers' emphasis with low-achieving students and all students in the same year.

Table 12. Total exposure to student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain

Teacher Report	Reading Test or Subtest—Bottom-quarter Students				Reading Test Subtest or Subtest—Top Three-quarters Students			
	Open-ended Reading	Closed-ended Reading			Open-ended Reading	Closed-ended Reading		
		Overall	Comprehension	Vocabulary		Overall	Comprehension	Vocabulary
Total Exposure (Frequency x Time Per Lesson) for Students:								
Read Materials of at Least One Paragraph		+		+				
Read Aloud		+		+	+			
Read Books They Choose Themselves			-					
Practice Word Attack					-	-	-	
Read Content Area Materials		+	+		+			
Practice Phonics				-		-		
Complete Workbooks/Skill Sheets								
Talk in Small Groups About What They Have Read	+	+	+		+		+	
Write About What They Have Read	-			-		-		
Work at a Computer		+		+				

Table reads: There was a significant ($p<.05$) positive relationship between (1) total exposure, in a fourth-grade teacher's classroom, to reading materials of at least one paragraph and (2) the gains made by that teacher's bottom-quarter students on the closed-ended reading test.

The frequency with which teachers reported having their students participate in an activity was associated with more test score gains than was the duration of these activities (Tables 13 and 14).

Table 13. Frequency of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain

Teacher Report	**Reading Test or Subtest—Bottom-quarter Students**				**Reading Test or Subtest—Top Three-quarters Students**			
	Open-ended Reading	**Closed-ended Reading**			**Open-ended Reading**	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**		**Overall**	**Comprehension**	**Vocabulary**
Frequency of Activity for Students:								
Read Materials of at Least One Paragraph	+	+	+	+				+
Read Aloud		+						
Read Books They Choose Themselves								
Practice Word Attack								
Read Content Area Materials								
Practice Phonics								-
Complete Workbooks/ Skill Sheets		+		+	+			
Talk in Small Groups About What They Have Read	+			+				
Write About What They Have Read								
Work at a Computer		+						

Table reads: There was a significant (p<.05) positive relationship between (1) frequency, in a fourth-grade teacher's classroom, of reading materials of at least one paragraph and (2) the gains made by that teacher's bottom-quarter students on the closed-ended reading test.

Table 14. Duration of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Reading Test or Subtest—Bottom-quarter Students				Reading Test or Subtest—Top Three-quarters Students			
	Open-ended Reading	Closed-ended Reading			Open-ended Reading	Closed-ended Reading		
		Overall	Comprehension	Vocabulary		Overall	Comprehension	Vocabulary
Time Spent per Lesson in Activity for Students:								
Read Materials of at Least One Paragraph								
Read Aloud		+		+				
Read Books They Choose Themselves			-					
Practice Word Attack							-	
Read Content Area Materials		+		+				
Practice Phonics				-				
Complete Workbooks/ Skill Sheets				-				
Talk in Small Groups About What They Have Read								
Write About What They Have Read								
Work at a Computer								

Table reads: There was a significant ($p<.05$) negative relationship between (1) time per lesson, in a fourth-grade teacher's classroom, in reading self-chosen books and (2) the gains made by that teacher's bottom-quarter students on the open-ended reading test.

The most frequently used activities in 1998 included reading materials of a paragraph or longer (90 percent with all students) and reading aloud (79 percent with all students) "almost every day." Teachers less frequently reported having students write about something they had read (32 percent with all students), complete reading workbooks or skill-sheet assignments (36 percent with all students), and work at a computer (29 percent with all students), as illustrated in Table 15.

Table 15. Frequency of student instructional activities in reading/language arts: percentage of fourth-grade teachers who reported having students engage in instructional activities by frequency

Instructional Activity	1998 Frequency (N=242)					1997 Frequency (N=218)				
	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Once or Twice a Semester	Never	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Once or Twice a Semester	Never
Read Materials of at Least One Paragraph	90	9	0	0	1	91	8	0	0	0
Read Aloud	79+	18+	3	0	0	68	30	1	0	0
Read Books They Choose Themselves	58	28	9	1	4	53	31	11	3	3
Practice Word Attack	52	28	10	3	7	44	34	13	5	5
Read Content Area Materials	46+	35	12+	2	5	56	35	5	2	2
Practice Phonics	40+	27	15	4	14	29	29	22	7	14
Complete Workbooks/Skill Sheets	36	38	16	3	6	35	41	13	3	7
Talk in Small Groups About What They Have Read	35	45	13+	3	4	27	42	23	4	4
Write About What They Have Read	32	44	20	1	3	29	48	18	3	2
Work at a Computer	29	52	9	3	7	28	48	12	5	8

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p < .05$ level) difference between 1997 and 1998.

Some differences did exist, however, between the frequency of instructional activities engaged in by all students and by low-achieving students. Specifically, statistically significant differences (those with p<.05) were found between activities conducted "almost every day" by all students and low-achieving students in the following areas:

- Practice phonics. Teachers more frequently assigned practice in phonics for low-achieving students (52 percent of teachers) than for all students (40 percent).
- Practice word attack. Sixty-two percent of teachers stated that they focus on this activity "almost every day" with low-achieving students. Only 52 percent reported this level of frequency with all of their students.
- Read materials of a paragraph or longer. This was an activity which was engaged in more frequently by all students (90 percent of teachers) than by low achievers (81 percent).

These results are not surprising, although they do not square with the findings on student achievement that have emerged in this study. Teachers may more frequently use early-reading activities such as practicing phonics and word attack with low-achieving students because they believe the students need this type of early literacy skill building. However, the findings of this study do not show a positive association between frequent practice in phonics and achievement gains for low achievers. It is also not surprising that all students (who, as a whole, have better reading skills) participate in more grade-level appropriate work such as reading materials of a paragraph or longer—yet more frequent reading of longer materials was more clearly associated with gains for the bottom-quarter students.

Of the changes in student instructional activities observed across the 2 years, many occurred in areas that had either no association or a negative association with growth in student achievement. For both groups of students, teachers increased the frequency of practicing phonics and decreased the frequency of reading in the content areas between 1997 and 1998, even though neither of these activities was associated with test score gains for either group. In fact, one activity that increased, practicing phonics, was negatively associated with student achievement for the top three-quarters of students on one test. What is appropriate for early learners, it appears, may not work well in fourth grade. This is offset somewhat by the statistically significant increase (p<.05) in the percentage of teachers reporting that they frequently have low-achieving students read aloud and talk in small groups about what they read—activities that were associated with test score gains for students in the bottom quarter of their class (Tables 11, 13, and 15).

Longer duration for a specific instructional activity in reading was negatively associated with student achievement growth in some cases. This negative relationship was found in many more cases for bottom-quarter students than top three-quarters students (Table 14).

There were some differences in the duration of students' instructional activities, either between groups of students or across years. For students of varying achievement levels, statistically significant differences ($p<.05$) were found between the percentage of teachers who reported reading materials of at least one paragraph (18 percent with all students, 31 percent with low-achieving students) and reading content area materials (25 percent with all students, 41 percent with low-achieving students) for "more than half a lesson." This suggests that teachers allow students of low academic achievement more time to read challenging material.

Between 1997 and 1998, several instructional activities tended to shorten in duration, although some of these were activities that increased in frequency (Tables 11 and 16). Fewer teachers reported that for "more than half a lesson" they had their students: read aloud (50 percent for all students in 1997, 34 percent in 1998), read content area materials (38 percent for all students in 1997, 25 percent in 1998); and practice phonics (70 percent for all students in 1997, 59 percent in 1998). This indicates that teachers are having students read aloud and practice phonics more often but for shorter periods of time.

Reading content area materials, however, seems to be a strategy that is used less (in both frequency and duration) in 1998 than in 1997. This is somewhat disappointing because some positive effects were associated with students' total exposure to this activity. Negative associations between the frequency or duration of an activity and student test scores may not indicate that a lesson is not beneficial in and of itself. Rather, this may simply indicate that the way in which an activity is approached (either how long students spend on it at a time, or how often they return to it) should be changed.

3.3 Work with Students of Varying Ability

The approaches that teachers used to work with students of different achievement levels were not associated with test score gains for any group of students. In fact, both heterogeneous grouping and homogeneous grouping were negatively associated with gains for bottom-quarter students. Homogeneous grouping was also negatively associated with growth on the open-ended reading test for top three-quarters students (Table 17).

Teachers most frequently reported relying on the following strategies when working with students of different achievement levels: giving extra time to low performers (58 percent responding "to a great extent"), using different instructional materials (53 percent), and using frequent assessments of performance (52 percent). Increases for these last two strategies (using different instructional materials and frequent assessments) were statistically significant (Table 18). Again, while it is somewhat encouraging that teachers use strategies that are negatively associated with student learning less frequently than other activities, it does not appear that any of these techniques have much of an effect.

Table 16. Duration of student instructional activities in reading/language arts: percentage of fourth-grade teachers who reported having students engage in instructional activities

Instructional Activity	**1998 Duration (N=208)**			**1997 Duration (N=195)**		
	> ½ a lesson	About ½ a lesson	< ½ a lesson	> ½ a lesson	About ½ a lesson	< ½ a lesson
Read Materials of at Least One Paragraph	18	49	33	26	45	29
Read Aloud	34+	41+	25	50	31	19
Read Books They Choose Themselves	52	28	21	55	30	15
Practice Word Attack	57	29	13	65	28	8
Read Content Area Materials	25+	51	24	38	46	17
Practice Phonics	59+	27	15	70	22	8
Complete Workbooks/ Skill Sheets	50	36	13	53	38	9
Talk in Small Groups About What They Have Read	53	31	16	52	36	13
Write About What They Have Read	35	40	25	36	41	23
Work at a Computer	39	33	27	40	32	28

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

Table 17. Instructional strategies in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Reading Test or Subtest—Bottom-quarter Students**				**Reading Test or Subtest—Top Three-quarters Students**			
	Open-ended Reading	**Closed-ended Reading**			**Open-ended Reading**	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**		**Overall**	**Comprehension**	**Vocabulary**
Extent of Use, in Teaching Students of Different Achievement Levels, of:								
Extra Time with Low Performers								
Different Instructional Materials								
Frequent Assessments								
Heterogeneous Grouping		-						
Homogeneous Grouping			-		-			
One-on-One Instruction								

Table reads: There was a significant (p<.05) negative relationship between (1) a fourth-grade teacher's self-reported use of homogeneous grouping and (2) the gains made by that teacher's bottom-quarter students on the comprehension subtest of the closed-ended reading test.

Table 18. Instructional strategies in reading/language arts: percentage of fourth-grade teachers who report use of instructional strategies with students of varying ability

Instructional Activity	**Extent Used, 1998 (N=241)**				**Extent Used, 1997 (N=219)**			
	Great extent	Moderate extent	Small extent	Not at all	Great extent	Moderate extent	Small extent	Not at all
Extra Time with Low Performers	58	35	7	0	51	36	12	1
Different Instructional Materials	53+	32+	13	2	43+	42+	12	3
Frequent Assessments	52+	38	10	0	41+	47	11	0
Heterogeneous Grouping	41	36	19	4	34	39	24	3
Homogeneous Grouping	40	36	19	5	32	38	23	8
One-on-One Instruction	39	40	19	2	38	34	26	2

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

3.4 Teachers' Self-Reported Preparation in Reading/Language Arts

In contrast to the sparse findings that resulted from our investigation of particular classroom practices, teachers' own assessment of their skill in teaching reading did have a distinct pattern of positive relationships with student gains. All students, but especially those with low initial performance, appear to have been better off with teachers who had confidence in their own skills as reading teachers (Table 19). Teachers who felt well prepared to work with heterogeneous groups had students who made better gains on both reading tests; those who felt well prepared to use a variety of assessment strategies had students who gained more ground on the open-ended reading test. Indeed, for students with low initial performance, better growth was associated with having a teacher who felt well prepared with respect to *any* of the skills that we asked about.

Table 19. Teacher preparation in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain

Teacher Report	Reading Test or Subtest—Bottom-quarter Students				Reading Test or Subtest—Top Three-quarters Students			
	Open-ended Reading	Closed-ended Reading			Open-ended Reading	Closed-ended Reading		
		Overall	Comprehension	Vocabulary		Overall	Comprehension	Vocabulary
How Well Prepared To:								
Use Small Group Instruction	+							
Take Existing Skills Into Account	+							
Integrate Reading/ Language Arts with Content Areas	+							
Teach Heterogeneous Groups	+		+		+			+
Use a Variety of Assessment Strategies	+				+			

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's self-reported level of preparation to use small-group instruction and (2) the gains made by that teacher's bottom-quarter students on the open-ended reading test.

Most teachers were confident in their ability to teach reading in most areas, although in each area there was still a substantial fraction of teachers expressing less than complete self-confidence. Greater than two-thirds of teachers reported that they felt "very well prepared" to teach using small group instruction for reading/language arts (72 percent) and to take into account students' existing skill levels when planning curriculum and instruction (71 percent). Sixty-four percent of teachers each reported that they were prepared to integrate reading/ language arts into other content areas and to teach heterogeneous groups, and 58 percent stated that they were well prepared to use a variety of assessment strategies. The level of preparation also seemed to be increasing. Between 1997 and 1998, statistically significant increases were found for teachers who stated they were "very well prepared" to take existing skills into account and to use a variety of assessment strategies (Table 20).

Table 20. Teacher preparation in reading/language arts: percentage of fourth-grade teachers who report their level of preparation to use a variety of instructional strategies

Teaching Strategy	Level of Preparation, 1998 (N=247)				Level of Preparation, 1997 (N=220)			
	Very Well Prepared	Fairly Well Prepared	Somewhat Prepared	Not Well Prepared	Very Well Prepared	Fairly Well Prepared	Somewhat Prepared	Not Well Prepared
Use Small Group Instruction	72	21+	6	0	64	30	6	0
Take Existing Skills into Account	71+	24	4	0	61	32	6	0
Integrate Reading/ Language Arts with Content Areas	64	31	4	0	62	31	7	1
Teach Heterogeneous Groups	64	28	5	2	57	36	7	0
Use a Variety of Assessment Strategies	58+	33+	9	0	48	43	7	2

Note: Rows may not total to 100 percent due to rounding.
+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

3.5 Teachers' Response to Standards-Based Reform in Reading/Language Arts

In reading, the fourth-grade teachers who considered themselves most closely attuned to state or local standards-based reform had students with test score gains that were neither better nor worse than those of their peers on most measures. In fact, only one statistically significant relationship was found between test gains and teachers' familiarity with four policy instruments (content standards, curriculum frameworks, state or district student assessments, and performance standards), or the extent to which each of those instruments was reflected in their own classroom curriculum: gains in comprehension scores among students in the top three-quarters of their class were positively associated with the extent that student assessments were reflected in their teachers' curriculum (Table 21). This lack of association may reflect, in part, the differences between skills measured on the Stanford 9 tests and the skills emphasized in state standards and state assessments; it is possible that student gains on the state tests (if those tests were administered every year) would in fact show an association with teachers' adherence to state standards, frameworks, and assessments.

Table 21. Policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain

Teacher Report	Reading Test or Subtest—Bottom-quarter Students				Reading Test or Subtest—Top Three-quarters Students			
	Open-ended Reading	Closed-ended Reading			Open-ended Reading	Closed-ended Reading		
		Overall	Comprehension	Vocabulary		Overall	Comprehension	Vocabulary
Teacher's Familiarity With:								
Student Assessments								
Performance Standards								
Content Standards								
Curriculum Frameworks								
Extent Reflected in Curriculum:								
Student Assessments							+	
Performance Standards								
Content Standards								
Curriculum Frameworks								

Table reads: There was a significant ($p<.05$) positive relationship between (1) the extent to which a fourth-grade teacher's curriculum reflected state or district student assessments and (2) the gains made by that teacher's top three-fourths students on the comprehension subtest of the closed-ended reading test.

In general, teachers rated themselves as quite familiar with state standards and assessments and asserted that they were incorporating these policy instruments into their classroom curriculum. In fact, greater than 80 percent of fourth-grade teachers reported that they were "very" or "moderately" familiar with each of the four policy instruments and that their reading/language arts curriculum reflected them to a "great" or "moderate" extent (Table 22). And, generally speaking, teachers appeared to be becoming more familiar with the policy instruments and to be implementing them in their instruction to a greater extent. The percentage of teachers reporting that they were familiar with and that their curriculum reflected the policy instruments increased from 1997 to 1998 in all but one area—the degree to which their curriculum reflects their state or district curriculum frameworks.

Table 22. Policy instruments in reading/language arts: percentage of teachers who are familiar with and implementing standards-based reforms in their classrooms

Policy Instrument	Fourth-Grade Teachers				All Teachers			
	1998 (N=244)		1997 (N=218)		1998 (N=1069)		1997 (N=1130)	
Familiarity With:	Very familiar	Moderately familiar	Very familiar	Moderately familiar	Very familiar	Moderately familiar	Very familiar	Moderately familiar
Student Assessments	57	33	55	33	56	37	54	37
Performance Standards	47	39	40	44	48+	42	42+	43
Content Standards	43	43	39	44	45	44	41	45
Curriculum Frameworks	39	44	37	47	41	47	40	43
Curriculum Reflects:	Great extent	Moderate extent	Great extent	Moderate extent	Great extent	Moderate extent	Great extent	Moderate extent
Student Assessments	44	42+	38*	56+*	49	42+	46*	47+*
Performance Standards	42	42+	36	53+	45+	44+	41+	49+
Content Standards	43	46+	36*	57+*	51+	41+*	45+*	48+
Curriculum Frameworks	39	47	39	51	44	47	43	47

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

* Indicates a statistically significant (at the p<.05 level) difference between fourth-grade teachers and all teachers in the same year.

Mostly negative relationships were found between changes in student achievement and the teacher's perception of the appropriateness of the policy instruments. That is, students tended to gain less with teachers who believe that content standards, student assessments, performance standards, and integration with the content areas are appropriate for their students (Table 23). The only exception to this finding was that students in the bottom quarter of their class gained more ground on the comprehension section of the Stanford 9 when their teachers felt that their curriculum frameworks were appropriate for their students.

There were no significant changes between 1997 and 1998 in the extent to which teachers rated standards, frameworks, and assessments as appropriate for their students. Most teachers continued to respond that they were "very" or "fairly" appropriate.

Table 23. Perceived appropriateness of policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Reading Test or Subtest—Bottom-quarter Students**				**Reading Test or Subtest—Top Three-quarters Students**			
	Open-ended Reading	**Closed-ended Reading**			**Open-ended Reading**	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**		**Overall**	**Comprehension**	**Vocabulary**
How Appropriate for the Teacher's Students Are:								
Content Standards							-	
Curriculum Frameworks			+					
Student Assessments				-				
Performance Standards							-	
Integration with Content Areas				-				

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's rating of the appropriateness of curriculum frameworks for his or her students and (2) the gains made by that teacher's bottom-quarter students on the comprehension subtest of the closed-ended reading test.

3.6 Families and Schools

In addition to questions about their classroom curriculum and instruction, teachers also answered questions about the role of families in the academic success of their children. The survey asked about the involvement of parents or guardians in their children's education and the extent to which children arrived at school "ready to learn."

For the bottom quarter of students, there was a positive relationship between teachers' reports on both parent involvement and students coming to school ready to learn, and fourth-grade student gains on the reading closed-ended test (Table 24). Unfortunately, only 3 percent of fourth-grade teachers reported moderate involvement of "most" parents of their low-achieving students, and 7 percent of fourth-grade teachers reported that "most" of their low-achieving students usually come to school prepared to learn (Table 25). One caveat is that teachers' rating of these variables may have been affected by how well students were doing in their class.

Table 24. Parent involvement and students ready to learn (reading): relationship between fourth-grade students' gains

Teacher Report	Reading Test or Subtest—Bottom-quarter Students				Reading Test or Subtest—Top Three-quarters Students			
	Open-ended Reading	Closed-ended Reading			Open-ended Reading	Closed-ended Reading		
		Overall	Comprehension	Vocabulary		Overall	Comprehension	Vocabulary
Teacher's Report on:								
For how many of your low-achieving students are parents or guardians at least moderately involved in school activities?		+	+					-
How many of your low-achieving students usually come to school prepared to learn?		+	+	+		+	+	

Table reads: There was a significant (p≤.05) positive relationship between (1) fourth-grade teacher's response on parent involvement for his or her low-achieving students and (2) the gains made by that teacher's bottom-quarter students on vocabulary subtest of the closed-ended reading test.

Table 25. Percentage distribution of responses for fourth-grade teachers and all teachers to parent involvement and students ready to learn survey items (reading)

Teacher Report	Fourth-grade Teachers 1998 (N=234)				All Teachers 1998 (N=1050)			
	Most	**Many**	**Some**	**Few/ None**	**Most**	**Many**	**Some**	**Few/ None**
For how many of your low-achieving students are parents or guardians at least moderately involved in school activities?	3	6	33	58	5	7	32	55
How many of your low-achieving students usually come to school prepared to learn?	7	14	55	24	8	15	48	30

In reading, a few of the fourth-grade teachers' survey responses did show relationships with student growth:

- Students whose achievement was initially low in relation to their classmates' achievement tended to gain more with those teachers who had them read materials of at least a paragraph, read materials in the content areas, and talk in small groups about what they had read.
- Returning frequently to a variety of instructional activities tended to have a positive association with student growth—and, by the same token, activities of long duration tended to be negatively associated with growth.
- Neither the skill emphases reported by fourth-grade teachers nor the strategies they try to use in working with students of different achievement levels showed much relationship with student growth.
- Teachers' belief that they were well prepared to use a variety of instructional techniques had a positive association with growth for their low-achieving students. Their self-reported level of preparation to work with heterogeneous groups and to use a variety of assessment strategies showed clear positive associations with growth for all students.
- Teachers' reported familiarity with the policy instruments of standards-based reform and the extent to which they believe they are following these policies in their classroom showed scant relationships to students' rates of growth.

Discouragingly, the kinds of practices that were associated with better student growth in reading were not necessarily the dominant ones in this sample of classrooms, and there was little systematic change in the direction of those practices between the 2 years of data collection.

4. MATHEMATICS CURRICULUM AND INSTRUCTION

This study has investigated the relationship between student growth in fourth-grade mathematics and the survey responses of that student's mathematics teacher with regard to curriculum and instruction. The analyses presented here generally parallel those just presented on the subject of reading curriculum and instruction. Again, we emphasize the association between teacher variables and rates of individual student growth in the fourth grade. We give particular attention to differences within classrooms, contrasting the students who had prior low achievement with the rest of the class, and we also note differences in teacher responses across the 2 years of the study. Data for all teachers in the study are generally not included here but instead are provided in the next chapters of the report.

In mathematics, just as in reading, we tested the hypothesis that teaching practice has no impact on scores and summarize the results in the tables that follow. The table format is again as follows: *If the hypothesis was not rejected* at the .05 significance level, the *cells are blank* in the tables.[6] If the *test was significant* at the .05 significance level, *either plus signs or minus signs* are shown in the tables. A plus sign indicates that increasing the quantity of an instructional method is related to a significant increase in test scores. Significant negative relationships are indicated by minus signs.

Most of the analyses presented here were performed separately for each of two groups of students in each teacher's classroom—those whose *third-grade* scores on the test or subtest fell into the bottom quarter, *relative to the scores of their own fourth-grade classmates on that test or subtest,* and those whose scores fell into the top three-quarters for their fourth-grade class. This analytic division of the "bottom quarter" and "top three-quarters" in each class gave us separate information about students who corresponded to two groups that were singled out in many of the survey questions for teachers, "your lowest achieving students (roughly the bottom quarter)" and "your top three-quarters students." This procedure allowed us to examine how teachers differ in their approach to dealing with students of varying skill levels and how curriculum and instructional techniques are working for different groups of students. A different analysis, included in Appendix B, looks at the relationship between teacher behavior and student gains for those students who are in the "bottom quarter" of *all fourth graders nationally*.

---

[6] With the large number of tests carried out here, some significant results will be purely due to chance. However, the overall rate of significant effects for math was 29 percent, which is significantly above the rate expected due to chance. Unless a teaching activity was statistically significant for a majority of the reading or math tests, we were skeptical of its effectiveness.

The areas in which we explored possible relationships to student rates of growth again include curriculum emphasis, the frequency and duration of instructional activities, the teacher's sense of preparedness in specific pedagogical skills, elements of standards-based reform, and family involvement in children's learning.

4.1 Topics and Skills Emphasized

Several survey items asked teachers about the number of lessons that they taught on particular topics during the year and, for each topic, the types of skills that they wanted students to learn. The skills, or cognitive demand, associated with the lesson were further broken down to distinguish between top three-quarters and low-achieving students—in other words, teachers were asked what they wanted their top three-quarters students to learn about a topic and also what they wanted their low-achieving students to learn.

First, we examine the concepts that teachers address with their students and the degree to which (or the number of lessons in which) teachers reported focusing on that topic over the course of the year. Then, we compare the skills that teachers emphasize with both "top three-quarters" and "low-achieving" students.

Positive relationships were found between the number of lessons taught in a particular subject and student gains on various sections of the Stanford 9, but more of these relationships were found for top three-quarters students (or those in the top three-quarters of their class) than for low-achieving students (Table 26). However, Table 27 shows that there was not a significant increase in the number of lessons taught in these activities between 1997 and 1998.

We also examined the emphasis placed on particular mathematical skills. Teachers were asked about the emphasis they placed on the following skills for all students and for low-achieving ones: memorizing facts, understanding concepts, solving equations, collecting/ interpreting data, solving word problems, and solving novel problems.

Skills that could be classified as critical thinking, rather than memorization, were more frequently associated with test score gains. A curriculum that focused on the skills of understanding concepts, solving equations, and solving problems seemed to be especially beneficial to students in the bottom quarter of their class, whose gains on the problem-solving subtest of the Stanford 9 were positively associated with their teachers' reported emphasis on these more cognitively demanding skills (Table 27).

Table 26. Topical coverage in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Mathematics Test or Subtest—Bottom-quarter Students**				**Mathematics Test or Subtest—Top Three-quarters Students**			
	Open-ended Mathematics	**Closed-ended Mathematics**			**Open-ended Mathematics**	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**		**Overall**	**Problem Solving**	**Procedures**
Number of Lessons Taught in:								
Word Problems with Addition, Subtraction					-			
Multi-Digit Multiplication						+		+
Rounding		+	+			+	+	+
Using Number Lines and Rulers	+	+	+			+		+
Operations with Fractions						+	+	+
Finding Length, Perimeter with Pictures						+		+
Solving Equations with One Unknown		+	+			+	+	
Distance Problems			+			+	+	+
Determining Central Tendency	+				+		+	-
Solving Equations with Two Unknowns						+	+	

Table reads: There was a significant ($p<.05$) positive relationship between (1) the number of lessons a fourth-grade teacher reported teaching on using number lines and rulers and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

Table 27. Cognitive demand in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Mathematics Test or Subtest—Bottom-quarter Students				Mathematics Test or Subtest—Top Three-quarters Students			
	Open-ended Mathematics	Closed-ended Mathematics			Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures		Overall	Problem Solving	Procedures
	Average Emphasis (Across Topics) on Teaching Low-Achieving Students to:				Average Emphasis (Across Topics) on Teaching Students to:			
Understand Concepts			+					-
Solve Equations			+					
Solve Word Problems			+					
Collect/ Interpret Data								
Memorize Facts					-			
Solve Novel Problems			+			+		+

Table reads: There was a significant ($p < .05$) positive relationship between (1) the average emphasis reported by a fourth-grade teacher on teaching students to understand concepts and (2) the gains made by that teacher's bottom-quarter students on the problem-solving subtest of the closed-ended mathematics test.

Top three-quarters students did not show as many positive associations between achievement gains and the cognitive demands of their curriculum. Although a positive relationship was found between test gains and the emphasis placed on solving novel problems, no relationship was found with solving equations or word problems for this group. Additionally, there was a negative relationship for top three-quarters students between scores on the procedures subtest and teachers' reported focus on understanding concepts.

Ironically, two of the skill emphases that were associated with better gains for low-achieving students were in fact reported significantly *less* often for such students. Teachers were less likely to emphasize understanding concepts and solving word problems with their low-achieving students (Table 28).

The skills emphasized by teachers changed to some extent between 1997 and 1998. Two areas of increasing emphasis were solving equations (for low-achieving students) and solving novel problems (for both groups of students). In fact, the number of teachers who indicated that they placed *no* emphasis on solving novel problems with low-achieving students dropped by 15 percentage points between 1997 and 1998 (Table 28). These changes in curriculum emphasis seem to be beneficial because solving novel problems was associated with gains for both groups of students, and solving equations was related to gains for bottom-quarter students.

4.2 Teachers' Instructional Activities

As for reading, we report on the gains associated with total exposure to teachers' instructional activities for particular groups of students, then break out total exposure into its components of frequency and duration (Tables 29-31). Student growth was positively associated with high levels of total exposure to activities that necessitated active student participation—such as taking a test, using manipulatives, and discussing multiple approaches to solving a problem—rather than those that were more teacher-focused—such as lecturing, presenting material using a blackboard, and teacher-led whole group discussion. It also appears, as for reading, that repeating an activity often was a good strategy but that remaining with a particular activity for a long time was not. Even teacher lectures, widely deplored by reformers, showed no negative relationship with achievement gain when they were frequent, although spending a high proportion of each lesson in lecturing was associated with lower rates of student growth.

The activities that teachers reported engaging in most frequently were associated with better rates of student achievement gain (Tables 30 and 32). These activities included working an exercise at the board (99 percent reported doing this at least once or twice a week in 1998), leading whole group discussions (94 percent), and discussing multiple approaches to problem solving (93 percent). Those activities that teachers reported conducting for the longest duration were associated with poorer rates of gain (Tables 31 and 33). These included lecturing (92 percent reported a duration of at least half a lesson in 1998) and working an exercise at the board (90 percent).

Table 28. Cognitive demand in mathematics: percentage of fourth-grade teachers who emphasized a skill

Subject	Degree of Emphasis in 1998 (N=203)				Degree of Emphasis in 1997 (N=194)			
	A Lot	Moderately	Occasionally	No Emphasis	A Lot	Moderately	Occasionally	No Emphasis
How much do you emphasize with all students?								
Understand Concepts	84*	15	1	0*	79	19	2	0*
Solve Equations	58	38	3*	0*	50	45	5*	1*
Solve Word Problems	44*	49	6*	1	41*	50	8*	1
Collect/ Interpret Data	37	50	11	1	32	55	10*	3
Memorize Facts	24	53	21	2	21	53	22	4
Solve Novel Problems	18	44	32	6+	14	37	29	20
How much do you emphasize with low-achieving students?								
Understand Concepts	76	22	1	1+	72	25	2	1
Solve Equations	50	38+	10	2	40	48	9	2
Solve Word Problems	35	51	12	1	28	54	16	2
Collect/ Interpret Data	30	52	15	2	26	50	21	4
Memorize Facts	23	49	23	6	22	46	26	6
Solve Novel Problems	17	40	32	10+	11	33	32	25

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

* Indicates a statistically significant (at the $p<.05$ level) difference between teachers' emphasis with low-achieving students and all students in the same year.

Table 29. Teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain

Teacher Report	**Mathematics Test or Subtest—Bottom-quarter Students**				**Mathematics Test or Subtest—Top Three-quarters Students**			
	Open-ended Mathematics	**Closed-ended Mathematics**			**Open-ended Mathematics**	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**		**Overall**	**Problem Solving**	**Procedures**
Total Exposure (Frequency x Time Per Lesson) to Teacher Activity:								
Work an Exercise at the Board	-				-		-	+
Lead Whole-Group Discussions					-			
Lecture or Present					-			
Discuss Multiple Approaches To Solving a Problem		+	+			+		+
Use Manipulatives	+							
Administer a Test		+	+	+		+		+

Table reads: There was a significant (p<.05) negative relationship between (1) total exposure reported by a fourth-grade teacher to the teacher working an exercise at the board and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

Table 30. Frequency of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Mathematics Test or Subtest—Bottom-quarter Students**				**Mathematics Test or Subtest—Top Three-quarters Students**			
	Open-ended Mathematics	**Closed-ended Mathematics**			**Open-ended Mathematics**	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**		**Overall**	**Problem Solving**	**Procedures**
Frequency of Teacher Activity:								
Work an Exercise at the Board		+	+	+		+		+
Lead Whole-Group Discussions	+		+	+				
Lecture or Present								
Discuss Multiple Approaches To Solving a Problem			+			+		+
Use Manipulatives	+							
Administer a Test				+		+		+

Table reads: There was a significant ($p<.05$) positive relationship between (1) frequency with which a fourth-grade teacher reported leading whole-group discussions and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

Table 31. Duration of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Mathematics Test or Subtest—Bottom-quarter Students				Mathematics Test or Subtest—Top Three-quarters Students			
	Open-ended Mathematics	Closed-ended Mathematics			Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures		Overall	Problem Solving	Procedures
Time per Lesson Spent in Teacher Activity:								
Work an Exercise at the Board	-				-		-	
Lead Whole-Group Discussions								
Lecture or Present	-	-		-	-	-	-	-
Discuss Multiple Approaches To Solving a Problem	-							+
Use Manipulatives		-					-	
Administer a Test						+	+	

Table reads: There was a significant (p<.05) negative relationship between (1) time per lesson reported by a fourth-grade teacher in working an exercise at the board and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

Table 32. Frequency of teacher instructional activities in mathematics: percentage of fourth-grade teachers who reported teacher instructional activities

Instructional Activity	1998 Frequency (N=202)					1997 Frequency (N=192)				
	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Once or Twice a Semester	Never	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Once or Twice a Semester	Never
Work an Exercise at the Board	86	13	1	1	0	86	11	1	1	1
Lead Whole-Group Discussions	68	26+	3	2	2	74	18	4	2	3
Lecture or Present	64+	25	5	2	4+	53	24	5	2	16
Discuss Multiple Approaches To Solving a Problem	46	47	4+	1	1	51	38	11	1	0
Use Manipulatives	32	46	21	1	0	27	48	22	1	2
Administer a Test	2+	46	44	5	2	0	51	41	6	2

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

Table 33. Duration of teacher instructional activities in mathematics: percentage of fourth-grade teachers who reported instructional activities

Instructional Activity	1998 Duration (N=178)			1997 Duration (N=177)		
	> ½ a Lesson	About ½ a Lesson	< ½ a Lesson	> ½ a Lesson	About ½ a Lesson	< ½ a Lesson
Work an Exercise at the Board	48+	42+	10	63	29	8
Lead Whole-Group Discussions	50	39	12	51	35	15
Lecture or Present	62+	30+	8	77	18	5
Discuss Multiple Approaches To Solving a Problem	45	39	16	48	36	15
Use Manipulatives	20+	44	36	31	42	27
Administer a Test	19	42	39	21	40	40

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

Few changes were found in the instructional activities that teachers used between 1997 and 1998. Those changes, however, indicated that teachers may be increasing the frequency of activities while shortening their duration—coming back to them more often but for shorter amounts of time—an approach that might well prove effective, according to our analysis. Specifically, the frequency with which teachers worked an exercise at the board, lectured or presented, and used manipulatives either increased or stayed the same between 1997 and 1998, while statistically significant decreases ($p<.05$) were found in the duration of these activities (Tables 32 and 33)

4.3 Students' Instructional Activities

For students' activities as for teachers' instructional activities, student growth was more likely to be associated with participation in activities that were incorporated frequently for a short duration (i.e., less than half a period). Achievement growth showed just a few significant associations with the total exposure to activities but many positive associations with the frequency of activities and many negative associations with the duration of activities (Tables 34-36).

Table 34. Student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Mathematics Test or Subtest—Bottom-quarter Students				Mathematics Test or Subtest—Top Three-quarters Students			
	Open-ended Mathematics	Closed-ended Mathematics			Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures		Overall	Problem Solving	Procedures
Total Exposure (Frequency x Time Per Lesson) to Student Activity:								
Respond Orally to Questions	-				-			
Work Individually on Worksheets					-	+		
Work in Small Groups						-		
Discuss Solutions in Whole Group		-			-			
Drill on Computational Skills					-			
Participate in Student-Led Whole-Group Discussions	-				-			
Analysis With Tables and Graphs								
Use Calculators To Solve Problems								
Assignments Requiring More Than a Paragraph								
Work With Manipulatives								
Assignments Taking More Than a Week		+	+	+		+	+	
Review Completed Homework in Class		+		+	-			

Table reads: There was a significant ($p<.05$) negative relationship between (1) total exposure reported by a fourth-grade teacher to oral response to questions and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

Table 35. Frequency of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Mathematics Test or Subtest—Bottom-quarter Students				Mathematics Test or Subtest—Top Three-quarters Students			
	Open-ended Mathematics	Closed-ended Mathematics			Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures		Overall	Problem Solving	Procedures
Frequency of Student Activity:								
Respond Orally to Questions								
Work Individually on Worksheets	+			+		+		
Review Completed Homework in Class	+	+	+			+		+
Drill on Computational Skills	+	+	+	+		+		+
Work with Manipulatives	+				+			
Work in Small Groups								
Discuss Solutions in Whole Group	+	-		-				
Assignments Requiring More than a Paragraph			+					
Participate in Student-Led Whole-Group Discussions	-				-			
Analysis with Tables and Graphs	+							
Use Calculators To Solve Problems	+				+			+
Assignments Taking More Than a Week	+	+	+	+		+	+	+

Table reads: There was a significant ($p<.05$) positive relationship between (1) frequency with which a fourth-grade teacher reported that students worked individually on worksheets and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

Table 36. Duration of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Mathematics Test or Subtest—Bottom-quarter Students				Mathematics Test or Subtest—Top Three-quarters Students			
	Open-ended Mathematics	Closed-ended Mathematics			Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures		Overall	Problem Solving	Procedures
Time Spent per Lesson in Student Activity:								
Respond Orally to Questions	-				-			
Work Individually on Worksheets	-		-		-			
Review Completed Homework in Class					-		-	
Drill on Computational Skills	-							
Work with Manipulatives								
Work in Small Groups					-	-		-
Discuss Solutions in Whole Group	-				-		-	
Assignments Requiring More Than a Paragraph								
Participate in Student-Led Whole-Group Discussions					-			
Analysis with Tables and Graphs								
Use Calculators To Solve Problems					-			
Assignments Taking More Than a Week	-							-

Table reads: There was a significant (p<.05) negative relationship between (1) the time per lesson reported by a fourth-grade teacher for students' oral response to questions and (2) the gains made by that teacher's bottom-quarter students on the open-ended mathematics test.

For several activities, there was a positive association between the frequency of the activities and gains by bottom-quarter students on more than one section of the Stanford 9. These activities also showed positive associations, but to a smaller degree, for higher achieving students. For example, drill on computational skills was associated with gains on all sections of the Stanford 9 for students in the bottom quarter of their class but was only associated with gains in the closed-ended test and the procedures section of that test for their classmates. Low-achieving students were also more likely to show negative associations between growth in performance and student activities, such as discussing mathematics in a whole-group setting (Table 35).

In contrast to our findings about teachers' instructional activities, we did not find alignment between the frequency of student instructional activities and the apparent effectiveness of those activities. For example, mathematics assignments that take a week or more to complete were the least frequently reported activity but were associated with test score gains on all sections of the Stanford 9 for bottom-quarter students and all closed-ended sections for top three-quarters students (Tables 35 and 37).

The variety of student instructional activities did increase from 1 year to the next, however. Statistically significant ($p<.05$) decreases were found between 1997 and 1998 in the percentage of teachers who reported "never" using: assignments taking more than 1 week to complete (38 percent), participating in student-led whole group discussions (30 percent), and assignments requiring writing more than a paragraph (18 percent) (Table 37). Some of these trends were in activities associated with student gains, while others were not.

As with teachers' instructional activities, the duration of a student activity was negatively associated with gains in test scores (Table 36). This was true more frequently for top three-quarters than for low-achieving students, perhaps because teachers had to give low-achieving students more time to complete their work. The exception was that no negative associations were found for lessons that would most likely require at least half a period to complete—such as assignments that require writing a paragraph or more, analysis with tables and graphs, and working with manipulatives. There does not appear to be a relationship between the percentage of teachers who reported conducting a particular activity for half a period or more and the extent to which that activity was related to student growth (Table 38). The year-to-year comparisons show a great deal of stability in the duration of student instructional activities from 1 year to the next.

Table 37. Frequency of student instructional activities in mathematics: percentage of fourth-grade teachers who report having students engage in instructional activities

Instructional Activity	1998 Frequency (N=204)					1997 Frequency (N=192)				
	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Once or Twice a Semester	Never	Almost Every Day	Once or Twice a Week	Once or Twice a Month	Once or Twice a Semester	Never
Respond Orally to Questions	95	4	0	1	0	95	3	2	1	0
Work Individually on Worksheets	65	30	3	1	1	73	23	2	1	2
Review Completed Homework in Class	52	34	8	2	4	45	41	6	3	6
Drill on Computational Skills	45	41	11	2	1	51	36	7	2	4
Work with Manipulatives	27	42	26	4	1	26	44	25	3	2
Work in Small Groups	26	54	16	1	2	22	50	21	5	2
Discuss Solutions in Whole Group	21	42	27	6	4	19	40	29	7	5
Assignments Requiring More Than a Paragraph	14+	41+	28	7+	10+	6	29	27	19	18
Participate in Student-Led Whole Group Discussions	11	30	26	12	22+	11	24	24	11	30
Analysis with Tables and Graphs	7	36	44	9	5	7	36	45	9	3
Use Calculators To Solve Problems	7	34	35	13	10	9	33	36	15	7
Assignments Taking More Than a Week	4	9	30	33	25+	4	4	25	29	38

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

Table 38. Duration of student instructional activities in mathematics: percentage of fourth-grade teachers who reported having students engage in instructional activities

Instructional Activity	1998 Duration (N=184)			1997 Duration (N=182)		
	> ½ a Lesson	About ½ a Lesson	< ½ a Lesson	> ½ a Lesson	About ½ a Lesson	< ½ a Lesson
Respond Orally to Questions	47+	35	18+	67	26	7
Work Individually on Worksheets	42	45	13	46	38	16
Review Completed Homework in Class	75	19	6	76	17	8
Drill on Computational Skills	58	33	8	51	29	11
Work with Manipulatives	18	48	34	24	43	33
Work in Small Groups	24	45	31	27	47	26
Discuss Solutions in Whole Group	37	43	20	40	42	17
Assignments Requiring More Than a Paragraph	32	41	27	39	41	20
Participate in Student-Led Whole Group Discussions	51	35	14	52	34	14
Analysis with Tables and Graphs	36	46	17	39	40	21
Use Calculators To Solve Problems	41	40	19	44	35	21
Assignments Taking More Than a Week	23	27	50	25	26	50

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

4.4 Teachers' Self-Reported Preparation in Mathematics Teaching

As in reading, the question asking teachers to rate their own skills in mathematics teaching yielded positive relationships with student gains, although in mathematics the relationships emerged only for those students who began the year with low achievement. For such students, good growth on some test or subtest was associated with all but one of the skills that we asked about (managing a class using manipulatives). Skills that seemed especially valuable were those of teaching heterogeneous groups and taking students' existing skills into account (Table 39).

There was not a great degree of variation in teachers' self-reported level of preparation in any of these areas. More than half of the teachers rated themselves as "very well prepared" in every skill except that of integrating mathematics with other subject areas. Teachers' level of preparation did seem to be increasing, however. In six of the eight areas, significantly more teachers stated that they were "very well prepared" to implement a strategy in 1998 than in 1997 (Table 40).

4.5 Teachers' Response to Standards-Based Reform in Mathematics

Although there was not a strikingly consistent relationship between teachers' reported disposition toward standards-based reform and their students' gains in mathematics, there was more of a pattern here than we found in reading. Bottom-quarter students' gains were greater on some subtests in those classrooms where the teachers were familiar with the policy instruments and implementing them in their curriculum. Specifically, positive relationships were found between bottom-quarter students' gains on the procedures subtest and their teacher's familiarity with performance standards and student assessments. On the other hand, there were negative relationships between the gains made by top three-quarters students and their teachers' familiarity with curriculum frameworks, content standards, and student assessments. Both groups of students, however, gained more when their teachers indicated that their curriculum reflected most of the policy instruments. Having performance standards and NCTM standards reflected in the curriculum showed the most positive associations (Table 41).

Table 39. Teacher preparation in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Mathematics Test or Subtest—Bottom-quarter Students**				**Mathematics Test or Subtest—Top Three-quarters Students**			
	Open-ended Mathematics	**Closed-ended Mathematics**			**Open-ended Mathematics**	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**		**Overall**	**Problem Solving**	**Procedures**
How Well Prepared To:								
Present Mathematics Concepts			+					
Teach Heterogeneous Groups	+	+	+	+				
Manage a Class Using Manipulatives								
Use Cooperative Learning Groups	+							
Use the Textbook as a Resource	+							
Take Students' Existing Concepts Into Account	+	+	+	+				
Use a Variety of Assessment Strategies		+		+	-			
Integrate Math with Other Subject Areas		+	+		-		-	

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's self-reported level of preparation to present mathematics concepts and (2) the gains made by that teacher's bottom-quarter students on the problem-solving mathematics subtest.

Table 40. Teacher preparation in mathematics: percentage of fourth-grade teachers who report their level of preparation to use a variety of instructional strategies

Teaching Strategy	Level of Preparation, 1998 (N=205)				Level of Preparation, 1997 (N=193)			
	Very Well Prepared	Fairly Well Prepared	Somewhat Prepared	Not Well Prepared	Very Well Prepared	Fairly Well Prepared	Somewhat Prepared	Not Well Prepared
Present Mathematics Concepts	66	31	3	0	57	38	4	1
Teach Heterogeneous Groups	66+	25+	9	0	46	45	8	1
Manage a Class Using Manipulatives	64	31	5	1	55	34	10	1
Use Cooperative Learning Groups	61+	31+	9	0+	43	43	11	3
Use the Textbook as a Resource	58+	33	9	0+	43	39	15	4
Take Students' Existing Concepts Into Account	55+	37	7	1	40	47	13	1
Use a Variety of Assessment Strategies	53+	39	7+	1	38	43	18	1
Integrate Math With Other Subject Areas	47+	36+	16	0	30	52	18	0

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

Table 41. Policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	Mathematics Test or Subtest—Bottom-quarter Students				Mathematics Test or Subtest—Top Three-quarters Students			
	Open-ended Mathematics	Closed-ended Mathematics			Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures		Overall	Problem Solving	Procedures
Teacher's Familiarity With:								
Student Assessments				+	-			
Performance Standards				+				
Content Standards					-		-	
Curriculum Frameworks					-	-	-	
Extent Reflected in Curriculum:								
Student Assessments		+						
Performance Standards		+		+	+			+
Content Standards			+					
Curriculum Frameworks								
NCTM Standards	+				+			

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's self-reported familiarity with state or district student assessments and (2) the gains made by that teacher's bottom-quarter of students on the procedures subtest of the closed-ended mathematics test.

Generally speaking, fourth-grade teachers reported an increasing familiarity with and implementation of standards-based reform between 1997 and 1998 (Table 42). Increases were reported in fourth-grade teachers' familiarity with student assessments and performance standards and in implementation of performance standards and NCTM standards. Most of these changes represented movement toward the teacher responses that were associated with better student gains.

Teachers in this sample were already quite familiar with standards-based reform in 1997, however. In fact, in 1997, greater than 80 percent of teachers already reported being "very or "moderately" familiar with standards and implementing them to a "great" or "moderate" extent. Specifically, teachers reported being "very" familiar with and implementing to a "great" extent student assessments (56 percent and 45 percent, respectively), performance standards (46 percent and 40 percent, respectively), content standards (41 percent and 44 percent, respectively), and curriculum frameworks (37 percent and 42 percent, respectively). This rate was lower for teachers who reported implementing NCTM standards—only 26 percent of teachers reported integrating these standards into their curriculum to a "great" extent in 1997 (Table 42).

As for reading, all but the lowest performing students showed better gains with those teachers who did not entirely believe their state or local standards-based reform framework was appropriate for the students they were teaching (Table 43). This relationship was found most frequently between test scores and teachers' responses to questions about performance standards and integration with the content areas. No statistically significant changes were found in the percentage of teachers who responded that the policy instruments were "very" or "fairly" appropriate between 1997 and 1998.

4.6 Families and Schools

In mathematics, the relationships between teachers' reports on both parent involvement and students coming to school ready to learn and fourth-grade student gains were not as clear as for reading. The bottom-quarter students had somewhat better gains on the closed-ended mathematics test if their teacher reported that they usually came to school prepared to learn (Table 44). As with reading, the percentage of teachers reporting that "most" parents of their low-achieving students are moderately involved in school activities and "most" students usually come to school ready to learn is low (Table 45). Two percent of fourth-grade teachers reported that "most" low-achieving students have parents who are at least moderately involved in school activities, and 7 percent of fourth-grade teachers reported that "most" of their low-achieving students usually come to school prepared to learn.

Table 42. Policy instruments in mathematics: percentage of teachers who are familiar with and implementing standards-based reforms in their classrooms

Policy Instrument	**Fourth-Grade Teachers**				**All Teachers**			
	1998 (N=202)		**1997 (N=194)**		**1998 (N=1007)**		**1997 (N=1074)**	
Familiarity With:	Very Familiar	Moderately Familiar	Very Familiar	Moderately Familiar	Very Familiar	Moderately Familiar	Very Familiar	Moderately Familiar
Student Assessments	56+	35	45	43	56+	36+	50	40
Performance Standards	46+	43	34	51	47+	43	41	45
Content Standards	41*	46	37	50	49+	41	44	44
Curriculum Frameworks	37	45	37	50	44	44	42	44
Curriculum Reflects:	Great Extent	Moderate Extent	Great Extent	Moderate Extent	Great Extent	Moderate Extent	Great Extent	Moderate Extent
Student Assessments	45	44	35*	53	48+	44	43	48
Performance Standards	40+	48+	27*	59	44+	47+	37	52
Content Standards	44	48	35*	57*	51+	43	46	47
Curriculum Frameworks	42	48	35*	55	47+	46	42	48
NCTM Standards	26+	38	17*	45	32+	40	26	40

Note: Rows may not total to 100 percent due to rounding.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.

* Indicates a statistically significant (at the $p<.05$ level) difference between teachers' emphasis with low-achieving students and all students in the same year.

Table 43. Perceived appropriateness of policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Mathematics Test or Subtest—Bottom-quarter Students**				**Mathematics Test or Subtest—Top Three-quarters Students**			
	Open-ended Mathematics	**Closed-ended Mathematics**			**Open-ended Mathematics**	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**		**Overall**	**Problem Solving**	**Procedures**
How Appropriate for the Teacher's Students Are:								
Student Assessments						-		-
Performance Standards					-	-	-	-
Content Standards						-		-
Curriculum Frameworks						-		-
Integration with Content Areas			+			-	-	-

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's rating of the appropriateness for his or her students of integrating mathematics with content areas and (2) the gains made by that teacher's bottom-quarter students on the problem-solving subtest of the closed-ended mathematics test.

Table 44. Parent involvement and students ready to learn (mathematics): relationship between fourth-grade teacher's response and fourth-grade students' gains

Teacher Report	**Mathematics Test or Subtest—Bottom-quarter Students**				**Mathematics Test or Subtest—Top Three-quarters Students**			
	Open-ended Mathematics	**Closed-ended Mathematics**			**Open-ended Mathematics**	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**		**Overall**	**Problem Solving**	**Procedures**
Teacher's Report on:								
For how many of your low-achieving students are parents or guardians at least moderately involved in school activities?			-	+				
How many of your low-achieving students usually come to school prepared to learn?		+		+				

Table reads: There was a significant ($p \leq .05$) positive relationship between (1) fourth-grade teacher's response on parent involvement for his or her low-achieving students and (2) the gains made by that teacher's bottom-quarter students on the procedures subtest of the closed-ended mathematics test.

Table 45. Percentage distribution of responses for fourth-grade teachers and all teachers to parent involvement and students ready to learn survey items (mathematics)

	Fourth-grade Teachers 1998 (N=197)				**All Teachers 1998 (N=997)**			
Teacher Report	**Most**	**Many**	**Some**	**Few/ None**	**Most**	**Many**	**Some**	**Few/ None**
For how many of your low-achieving students are parents or guardians at least moderately involved in school activities?	2	7	31	60	5	7	32	55
How many of your low-achieving students usually come to school prepared to learn?	7	11	55	27	8	14	47	30

4.7 Conclusions

This study has been more successful in finding teacher variables associated with student growth in fourth-grade mathematics than it has in fourth-grade reading. This may well reflect the state of the art in research on the two areas, with a base of prior research having enabled us to ask more discriminating questions about mathematics teaching. Alternatively, it is possible that there is more homogeneity in teacher quality with respect to reading, while skill in teaching mathematics varies more.

In any case, several kinds of teacher responses were associated with better rates of student growth in fourth-grade mathematics, especially for students who started out with the lowest achievement:

- Emphasis on the relatively demanding competencies such as problem solving and, for low-achieving students, conceptual understanding
- Frequent use of a wide variety of teacher and student instructional activities, especially those that require more active thinking by students. As in reading, however, devoting a high proportion of each lesson to a single activity was negatively associated with student growth for most activities
- The teacher's self-assessment as well prepared with respect to several specific skills in mathematics teaching
- The extent to which students arrived at school ready to learn

5. POLICY ENVIRONMENT AND PROFESSIONAL DEVELOPMENT: AN INITIAL EXPLORATION OF INFLUENCES ON TEACHERS

Ultimately, the LESCP study seeks to understand not only the relationships between classroom conditions and student outcomes in the sample schools but also the influences on these teachers' curriculum and instruction. For this interim report, we present the results of some early analyses of likely influences on teachers. We have explored some of the influences that may be found in Box 1c of the study's conceptual framework, the policies enacted by states and school districts. We have also looked at professional development, reasoning that it is one of the most salient aspects of Box 2, Implementation, as a possible influence on classroom curriculum and instruction.

First, this chapter looks at the policy environment around teachers, as reflected in documents that the study collected at the district level. We begin by describing our analysis of variation associated with the extent to which each of the 18 districts in the sample displayed a policy environment of standards-based reform—specifically, the extent of emphasis on standards, standards-based assessment, and accountability in formal statements of district policy. Based on documents obtained from districts, this analysis permitted us to explore differences in teachers' survey responses, especially with respect to the questions about standards-based reform, under different policy conditions.

Second, we discuss the study's findings on professional development, which we would expect to be one of the strongest avenues by which districts and schools could influence teachers' knowledge, skills, and behavior. This chapter describes the preparation and professional development reported by teachers in reading and mathematics, with particular attention to variation across policy environments. It then identifies relationships between teachers' professional development and the curriculum and instruction they reported for their classrooms. In this last area, the study has as yet found few relationships.

5.1 Policy Environments in the LESCP Districts

As described in Chapter 1 of this report, the 18 LESCP districts were classified according to the extent to which they had enacted policies on standards and aligned curriculum, assessment, and accountability. This analysis permitted us to identify 4 districts having the most clearly and thoroughly specified policies on these subjects, 4 districts that were comparatively lacking in such policies, and the other 10 districts in the middle.

First, we look at the distribution of all teachers' responses, across years, to the questions about their familiarity with standards, assessments, and curriculum frameworks and about their adherence to these policy instruments in the curriculum. There were upward trends in teachers' familiarity with a few policy instruments of standards-based reform in reading, as well as in the extent to which they said their curriculum reflected these policy instruments, between 1997 and 1998. As illustrated in Table 46, student assessments remained the most familiar policy instrument, but standards and frameworks were comparably important in their influence on classroom curriculum.

In mathematics, increases in familiarity with and implementation of these policy instruments were more pronounced. The percentage of teachers reporting that they are familiar with them and that they are incorporated into the curriculum increased in all but one area (familiarity with curriculum frameworks) between 1997 and 1998. As with reading, student assessments were the policy instruments with which teachers reported the greatest familiarity. Comparable percentages of teachers also reported that their curriculum reflected state or district-level standards, assessments, and curriculum frameworks to a "great extent" (Table 47).

When we break these totals down by categories of districts, according to the district policy environment with regard to standards-based reform, the results show differences in teachers' reports, in the expected direction. More teachers in districts with higher reform policy environments reported being familiar with and integrating the policy instruments into their curriculum than did so in the lower reform districts.

Table 46. Change in teacher familiarity with and adherence to policy instruments: reading

Policy Instrument in Reading	**1998 (N=1069)**	**1997 (N=1130)**
Percentage of Teachers "Very Familiar"		
Content Standards	45	41
Curriculum Frameworks	41	40
Student Assessments	56	54
Performance Standards	48+	42
Percentage of Teachers Whose Curriculum Reflects to a "Great Extent"		
Content Standards	51+	45
Curriculum Frameworks	44	43
Student Assessments	49	47
Performance Standards	45+	41

+ Indicates a statistically significant (at the $p < .05$ level) difference between 1997 and 1998.

Table reads: In 1997, 41% of all responding teachers reported they were "very familiar" with content standards in reading; in 1998, 45% did so. This change was not statistically significant.

Table 47. Change in teacher familiarity with and adherence to policy instruments: mathematics

Policy Instrument in Mathematics	**1998 (N=1009)**	**1997 (N=1076)**
Percentage of Teachers "Very Familiar"		
Content Standards	49+	44
Curriculum Frameworks	44	42
Student Assessments	56+	50
Performance Standards	47+	41
Percentage of Teachers Whose Curriculum Reflects to a "Great Extent"		
Content Standards	51+	46
Curriculum Frameworks	46	42
Student Assessments	48+	43
Performance Standards	44+	37
NCTM Standards	32+	26

+ Indicates a statistically significant (at the $p < .05$ level) difference between 1997 and 1998.

Table reads: In 1997, 44% of all responding teachers reported they were "very familiar" with content standards in mathematics; in 1998, 49% did so. This change was statistically significant.

There were few statistically significant changes ($p<.05$) in responses between 1997 and 1998 for any of the groupings we studied, but most changes between years occurred in "other" districts, or districts whose level of standards-based reform was somewhere in the middle of the range of districts that we studied. In reading, three of the five statistically significant ($p<.05$) increases were found with teacher responses from other districts. In math, all of the five changes between 1997 and 1998 occurred among teachers in other districts. Specifically, more teachers reported being "very" familiar with content standards (in reading and math), student assessments (in math), and performance standards (in reading and math). Increases in teacher responses that the curriculum reflects policy instruments to a "great" extent also occurred with this group of teachers for content standards in reading and student assessments, performance standards, and NCTM standards in math (Tables 48 and 49).

Table 48. Teachers' familiarity with policy instruments in reading: percentage of teachers familiar with and implementing policy instruments, by district policy environment

Policy Instrument	**1998**			**1997**		
	High-reform Districts (N=197)	**Low-reform Districts (N=98)**	**Other Districts (N=774)**	**High-reform Districts (N=205)**	**Low-reform Districts (N=105)**	**Other Districts (N=821)**
Percentage of Teachers "Very Familiar"						
Content Standards	51*	31^	46+	53*~	29^	40
Curriculum Frameworks	55*~	31	38	53*~	25^	39
Student Assessments	56*	40^	58	62*~	41^	53
Performance Standards	51*	29^	49+	51*~	27^	42
Percentage of Teachers Whose Curriculum Reflects to a "Great Extent"						
Content Standards	58+*	39^	50+	49*	32^	45
Curriculum Frameworks	52*~	38	42	49*	34	43
Student Assessments	51*	37^	50	54*	31^	47
Performance Standards	45	36+^	47	47*	22^	42

Table reads: In districts with high-reform policy environments (according to indicators described in the text above), 51 percent of teachers reported they were very familiar with content standards in reading in 1998.

+ Indicates a statistically significant (at the $p<.05$ level) difference between 1997 and 1998.
* Indicates a statistically significant (at the $p<.05$ level) difference between low-reform and high-reform districts.
~ Indicates a statistically significant (at the $p<.05$ level) difference between high-reform and other districts.
^ Indicates a statistically significant (at the $p<.05$ level) difference between low-reform and other districts.

Table 49. Teachers' familiarity with policy instruments in mathematics: percentage of teachers familiar with and implementing policy instruments, by district policy environment

Policy Instrument	1998			1997		
	High-reform Districts (N=185)	**Low-reform Districts (N=97)**	**Other Districts (N=727)**	**High-reform Districts (N=194)**	**Low-reform Districts (N=103)**	**Other Districts (N=780)**
Percentage of Teachers "Very Familiar"						
Content Standards	53*	40	49+	51	42	43
Curriculum Frameworks	55*~	35	42	50*~	32	41
Student Assessments	56	46^	57+	55*	40	50
Performance Standards	48	37^	48+	47*	31^	41
Percentage of Teachers Whose Curriculum Reflects to a "Great Extent"						
Content Standards	57	50	50	51	45	45
Curriculum Frameworks	55~	47	44	50*~	37	41
Student Assessments	48	47	48+	47	44	42
Performance Standards	46	46	44+	39	41	35
NCTM Standards	32	38	31+	27	28	25

Table reads: In districts with high-reform policy environments (according to indicators described in the text above), 53 percent of teachers reported they were very familiar with content standards in mathematics in 1998.

+ Indicates a statistically significant (at the p<. 05 level) difference between 1997 and 1998.
* Indicates a statistically significant (at the p<. 05 level) difference between low-reform and high-reform districts.
~ Indicates a statistically significant (at the p<. 05 level) difference between high-reform and other districts.
^ Indicates a statistically significant (at the p<. 05 level) difference between low-reform and other districts.

This seems to indicate that most growth in standards-based reform is happening in those districts that are engaged in standards-based reform to some degree but that have not yet reached full implementation. High-reform districts may be so advanced in their implementation of reform that little growth is possible, and it may be difficult to bring about change in low-reform districts. The factors that helped us identify the outliers on both ends of the range may inhibit our ability to find change.

We also examined differences between the groupings of districts—high reform, low reform, and other—to determine how pronounced the differences between them were. As would be expected, the greatest number of statistically significant ($p<.05$) differences in responses between groups of teachers were found between high- and low-reform districts for both reading (15 differences) and math (6 differences). However, there were also many differences between low-reform and other districts (16 differences total).

5.2 Professional Development and Preparation in Reading

The reauthorized Title I emphasizes professional development for teachers to ensure that they possess the knowledge and skills to help all children learn to high standards. It is reasonable to assume, also, that professional development organized around standards, assessments, and curriculum frameworks is an important vehicle for bringing teachers on board with standards-based reform. Thus, we investigated variation in teachers' reported preparation and their recent participation in professional development, with emphasis on the variation that might be associated with local policy environments.

As throughout the LESCP study, we examined professional development by reading and math. The teachers were asked about the following components of professional development:

- How well prepared they were to use a variety of instructional strategies
- The amount and quality of professional development received in content areas, instructional strategies, and parent involvement
- The extent to which the professional development was designed to support the policy environment
- The extent to which the professional development enhanced their knowledge and skills

There was no change in teachers' self-reported preparation in reading/language arts from 1997 to 1998. Teachers were most likely to describe themselves as "very well prepared" to use small group instruction for reading/language arts (74 percent in 1998 and 72 percent in 1997) and least likely to say this about using a variety of assessment strategies (55 percent in 1998 and 53 percent in 1997) (Table 50). As indicated previously in this report, fourth-grade teachers' self-reported preparation was positively related to their students' gains in reading, especially for those students with low initial performance.

Table 50. Percentage of teachers of reading who responded that they are "very well prepared" in selected teaching strategies

Teaching Strategies	1998			1997		
	All (N=1074)	High Reform (N=200)	Low Reform (N=97)	All (N=1032)	High Reform (N=203)	Low Reform (N=106)
Use small group instruction for reading/language arts	74	78	69	72	82	74
Take into account students' existing skills when planning curriculum and instruction	69	74	68	66	77+	60
Integrate reading/ language arts into other content areas	69	64	60	68	69	67
Use a variety of assessment strategies	55	55	57	53	58	51
Teach groups that are heterogeneous in ability	62	60	70	62	67	67

+ Indicates a statistically significant (at the .p<05 level) difference between high and low reform for 1997 or 1998.

The percent of teachers responding "very well prepared" to selected teaching strategies varied little between schools in high-reform districts and schools in low-reform districts, with little clustering of teachers by school. In 1997, teachers in high-reform districts were more likely than teachers in low-reform districts to report that they were "very well prepared" to take into account students' existing skills when planning curriculum and instruction (77 percent in high-reform districts and 60 percent in low-reform districts). However, this difference disappeared in 1998 (74 percent in high-reform districts and 68 percent in low-reform districts). For schools in the low-reform districts, the percentage of teachers responding "very well prepared" increased, though not significantly, for three of the five teaching strategies across the years (Table 50).

Although teachers in general reported less professional development on selected topics in 1998 than they reported in 1997, most teachers in the LESCP schools participated in some professional development. In general, this might indicate that the school or district emphasis on professional development has declined, or that the focus on professional development varies from year to year. Almost three-fourths of the teachers of reading reported participating in some professional development in content in reading, instructional strategies for teaching reading, and strategies for using assessment results. Slightly more than one-half of the teachers reported participation in professional development focused on instructional strategies for teaching low-achieving students (Table 51). Of those who had participated in any professional development in these topics, less than 50 percent rated the quality as high. Only 30 percent of teachers who participated in professional development focused on strategies for using assessment results rated the quality as high.

Table 51. Percentage of teachers of reading who responded that they participated in "no" professional development on selected topics

Topic Area	**1998**			**1997**		
	All (N=1042)	**High Reform (N=194)**	**Low Reform (N=95)**	**All (N=1088)**	**High Reform (N=194)**	**Low Reform (N=105)**
Content in reading	25+	49*	24	21	42*	17
Instructional strategies for teaching reading	25	54*	18	22	45*	28
Strategies for using assessment results	28+	29*	43	20	22*	34
Instructional strategies for teaching low-achieving students	41+	49	55	34	41*	61
Instructional strategies for teaching LEP students	73	91*	67	70	89*	74
Strategies to increase or strengthen parent involvement	56+	59+	62	41	44	54

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

* Indicates a statistically significant (at the p<.05 level) difference between high and low reform for 1997 or 1998.

In 1997 and 1998, teachers from high-reform districts most often reported professional development activities focused on strategies for using assessment results, while teachers in low-reform districts most often reported professional development activities focused on content in reading or instructional strategies for teaching reading. This may indicate that the high-reform districts are using assessment results for planning and continuous improvement of the instructional program, while low-reform districts are focused on the traditional types of professional development and limited use of assessment results for planning and improvement of the instructional program.

There was considerable variation across the districts in the amount and emphasis on professional development focused on reading. In 1998, more than 50 percent of the teachers in two districts reported "more than 2 days" of professional development focused on content in reading, and/or instructional strategies for teaching reading. In contrast, more than 50 percent of teachers in four districts reported no professional development focused on content in reading and more than 50 percent of teachers in three districts reported no professional development focused on instructional strategies for teaching reading. More than 50 percent of the teachers in three of the four districts reported no professional development in either category. In only one of the districts in the LESCP study did more than 50 percent of the teachers report "more than 2 days" of professional development in both content in reading and instructional strategies for teaching reading.

Professional development tended to be designed to support reform efforts at the school level in low-reform districts, while teachers in high-reform districts described their professional development as being focused on district or state reforms. Teachers in high-reform districts (34 percent) were more likely than teachers in low-reform districts (22 percent) to report that professional development activities were designed to support the state or district assessment to a "great extent." Teachers in low-reform districts were most likely to report that professional development activities were designed to support reform efforts under way in their school to a "great extent" (29 percent) (Table 52).

Teachers who participated in professional development activities valued the relevance of the activities. Forty-six percent of teachers gained confidence (ratings $\geq$ 4 on a 5-point scale) in using new pedagogical approaches in teaching reading/language arts as a result of professional development over the past year. Eighteen percent of the teachers said they had, to a "great extent," gained confidence (Table 53).

However, despite extensive exploration of the survey data, we found that teachers who reported gaining confidence in using new pedagogical approaches did not report high levels of any particular classroom practices. Similarly, they did not significantly increase their use of any particular practices across the 2 years of data collection.

Table 52. Percentage of teachers of reading who responded that their professional development activities were designed to support reform efforts to a "great extent"

Professional Development Activity	**1998**		
	All (N=1009)	**High Reform (N=178)**	**Low Reform (N=87)**
Well matched to your school's or department's plan to change practice	30	24	28
Designed to support reform efforts under way in your school	33	27	29
Designed to support state or district standards or curriculum frameworks	36	33	27
Designed to support state or district assessment	36	34*	22

*Indicates a significant difference at the p < .05 level between high- and low-reform for 1998.

Table 53. Percentage of teachers of reading who responded that their knowledge and skills were enhanced to a "great extent" as a result of professional development experiences during the past year

Professional Development Activity	**1998**		
	All (N=1011)	**High Reform (N=179)**	**Low Reform (N=87)**
Helped me adapt my teaching to meet state assessment requirements	23	18	15
Helped me adapt my teaching to meet state standards or curriculum framework requirements	22	20	15
Learned how to help students engage in collaborative inquiry	16	12	8
Gained confidence in using new pedagogical approaches in teaching reading/English/language arts	18	17	9
Feel more motivated to draw from a wide variety of methods when teaching	28	22	23

5.3 Professional Development and Preparation in Mathematics

Unlike the stability found in reading, teachers' reported level of preparation rose significantly with respect to three of the teaching strategies asked about in mathematics. The percent of teachers responding "very well prepared" significantly increased from 1997 to 1998 for the following teaching strategies: integrate mathematics with other subject areas (53 percent in 1998 and 48 percent in 1997), use a variety of assessment strategies (55 percent in 1998 and 49 percent in 1997), and teach groups that are heterogeneous in ability (63 percent in 1998 and 58 percent in 1997). Some differences (i.e., take into account students' prior conceptions about mathematics when planning curriculum and instruction, integrate mathematics with other subject areas, and use the textbook as a resource rather than as the primary instructional tool) existed between high- and low-reform districts in 1997, but the differences diminished in 1998 as low-reform districts came to look more like high-reform districts (Table 54).

Table 54. Percentage of teachers of math who responded that they are "very well prepared" in selected teaching strategies

Teaching Strategies	1998			1997		
	All (N=1010)	High Reform (N=184)	Low Reform (N=97)	All (N=1079)	High Reform (N=197)	Low Reform (N=105)
Present the applications of mathematics	66	66	63	65	72	71
Use cooperative learning groups in mathematics	56	59	55	53	57	47
Take into account students' prior conceptions about mathematics when planning curriculum and instruction	55	57	53	52	58*	40
Integrate mathematics with other subject areas	53+	54	46	48	56*	37
Manage a class of students who are using manipulatives	68	66	69	66	72	62
Use a variety of assessment strategies	55+	51	48	49	55	47
Use the textbook as a resource rather than as the primary instructional tool	61	56	55	57	60*	47
Teach groups that are heterogeneous in ability	63+	61	57	58	60	59

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

* Indicates a statistically significant (at the p<.05 level) difference between high and low reform for 1997 or 1998.

Overall, teachers were most likely to describe themselves as "very well prepared" to manage a class using manipulatives (68 percent in 1998 and 66 percent in 1997) and present the applications of mathematics concepts (66 percent in 1998 and 65 percent in 1997). They were least likely to say "very well prepared" about integrating math into other subject areas (53 percent in 1998 and 48 percent in 1997) and use a variety of assessment strategies (55 percent in 1998 and 49 percent in 1997) (Table 54). As previously reported, fourth-grade teachers' self-reported level of preparation was positively related to student growth for their low-achieving students.

As in reading, mathematics teachers, in general, reported less professional development on selected topics in 1998 than they reported in 1997. Although almost three-fourths of the teachers reported participating in professional development activities focused on content in mathematics and instructional strategies for teaching mathematics in 1998, this was a significant decrease from 1997. In 1997, approximately fourth-fifths of the teachers reported participating in professional development activities focused on content in mathematics or instructional strategies for teaching mathematics (Table 55). Of those who had participated in any professional development in these topics, less than 50 percent of the teachers rated the quality as high.

In 1998, teachers in high-reform districts were more likely than teachers in low-reform districts to participate in professional development activities that focused on content in mathematics or strategies for using assessment results. In both 1997 and 1998, teachers in low-reform districts were more likely to participate in professional development activities that focused on content in reading or instructional strategies for teaching reading.

As in reading, there was considerable variation across the districts in the amount and emphasis on professional development in mathematics. In 1998, more than 50 percent of the teachers in two districts reported "more than 2 days" of professional development focused on content in mathematics and instructional strategies for teaching mathematics. In contrast, more than 50 percent of teachers in five districts reported no professional development focused on content in mathematics, and in four districts, more than 50 percent of the teachers reported no professional development focused on instructional strategies for teaching mathematics. More than 50 percent of the teachers in three districts reported no professional development in either category.

Table 55. Percentage of teachers of math who responded that they participated in "no" professional development on selected topics

Topic Area	**1998**			**1997**		
	All (N=987)	**High Reform (N=180)**	**Low Reform (N=95)**	**All (N=1047)**	**High Reform (N=188)**	**Low Reform (N=103)**
Content in mathematics	27+	41*	58+	20	35	36
Instructional strategies for teaching mathematics	28+	48+	52	19	35	40
Strategies for using assessment results	29+	31*	45	21	25	35
Instructional strategies for teaching low-achieving students	43+	54+	57	35	41*	62
Instructional strategies for teaching LEP students	73+	94*+	69	69	88*	76
Strategies to increase or strengthen parent involvement	56+	59+	62	42	45	54

+ Indicates a statistically significant (at the p<.05 level) difference between 1997 and 1998.

* Indicates a significant difference (at the p<.05 level) between high and low reform for 1997 or 1998.

Professional development activities were most often designed to support reform efforts at the school level in low-reform districts, while teachers in high-reform districts more often said that professional development was focused on district or state reforms. Teachers in high-reform districts (36 percent) were more likely than teachers in low-reform districts (22 percent) to report that professional development activities were designed to support the state or district assessment to a "great extent." Teachers in low-reform districts were most likely to report that professional development activities were designed to support reform efforts under way in their school to a "great extent" (29 percent) (Table 56).

Table 56. Percentage of teachers of math who responded that their professional development activities supported reform efforts to a "great extent"

Professional Development Activity	1998		
	All (N=957)	**High Reform (N=165)**	**Low Reform (N=88)**
Well matched to your school's or department's plan to change practice	30	26	26
Designed to support reform efforts under way in your school	33	29	27
Designed to support state or district standards or curriculum frameworks	36	36	26
Designed to support state or district assessment	36	36*	22

* Indicates a statistically significant (at the $p<.05$ level) difference between high and low reform for 1998.

Teachers who participated in professional development gave cautiously favorable ratings to the relevance of the these activities. Forty-five percent of teachers gained confidence (ratings ≥ 4 on a 5-point scale) in using new pedagogical approaches in teaching math. Seventeen percent said they had, to a "great extent," gained confidence in using new pedagogical approaches in teaching math as a result of professional development over the past year. Teachers in high-reform districts (14 percent) were more likely than teachers in low-reform districts (6 percent) to report this enhancement to a "great extent" as a result of professional development. Referring to the extent to which professional development was designed to align with the policy environment, fewer than one-fourth of teachers in general (23 percent), as well as teachers in high-reform districts (22 percent) and low-reform districts (15 percent), reported that their knowledge or skills were enhanced to a "great extent" as a result of professional development that focused on adapting teaching to meet state standards or curriculum framework requirements (Table 57).

Table 57. Percentage of teachers of math who responded that their knowledge and skills were enhanced to a "great extent" as a result of professional development experiences during the past year

Professional Development Activity	1998		
	All (N=958)	**High Reform (N=166)**	**Low Reform (N=88)**
Helped me adapt my teaching to meet state assessment requirements	23	20	15
Helped me adapt my teaching to meet state standards or curriculum framework requirements	23	22	15
Learned how to help students engage in collaborative inquiry	16	12	8
Gained confidence in using new pedagogical approaches in teaching math	17	14*	6
Feel more motivated to draw from a wide variety of methods when teaching	29	23	20

* Indicates a statistically significant (at the p<.05 level) difference between high and low reform for 1998.

5.4 Relationship Between Professional Development and Teachers' Reports on Policy Instruments

For these analyses, we compared teachers' answers to the questions about policy instruments with the content of selected professional development activities. Content standards, curriculum frameworks, student assessments, and performance standards were the policy instruments used in these comparisons. The teachers were asked a variety of questions related to their participation in professional development. Specifically, teachers were asked:

1. Please indicate the amount of professional development you received in the past 12 months and if you received professional development, rate the quality.
    - Content in reading
    - Content in mathematics
    - Instructional strategies for teaching reading
    - Instructional strategies for teaching mathematics

- Strategies for using assessment results
- Instructional strategies for teaching low-achieving students
- Instructional strategies for teaching limited-English-proficient students
- Strategies to increase or strengthen parent involvement

2. To what extent was the professional development activity:
   - Well matched to your school's or department's plan to change practice?
   - Designed to support reform efforts under way in your school?
   - Designed to support state or district standards or curriculum frameworks?
   - Designed to support state or district assessment?

3. To what extent do you feel that your knowledge and skills have been enhanced in each of the following ways as a result of your participation in the professional development experiences you have had in the past year?
   - Helped me adapt my teaching to meet state assessment requirements
   - Helped me adapt my teaching to meet state standards or curriculum framework requirements
   - Learned how to help students engage in collaborative inquiry
   - Gained confidence in using new pedagogical approaches in teaching math
   - Gained confidence in using new pedagogical approaches in teaching reading/English/ language arts
   - Feel more motivated to draw from a wide variety of methods when teaching

In general, modest positive correlations (i.e., ≥ 0.25) were found between the policy instruments and the teachers' report of the extent to which the professional development was designed to support the policy environment. The findings are similar for both reading and math. Familiarity with content standards was modestly correlated with professional development activities designed to support state or district standards or curriculum frameworks (0.27), and professional development activities designed to support state or district assessments (0.26). All of the policy instruments were modestly correlated with professional development activities that helped teachers adapt their teaching to meet state assessment requirements or helped teachers to adapt their teaching to meet state standards or curriculum framework requirements (Table 58). Not unexpectedly, either professional development activities that addressed policy instruments increased the likelihood of teacher familiarity with these instruments, or those teachers who were more familiar with the policy instruments were more likely to recognize them as a focus of their professional development.

No correlations were found between the *amount* of professional development and familiarity with the policy instruments.

We also examined the relationship between the extent to which teachers believed that their curriculum reflected each of the policy instruments and the content of professional development. For this analysis, the NCTM standards were added to the mathematics section.

Again, only a few modest positive correlations (i.e., ≥ 0.25) were found between the reported integration of the policy instruments into the curriculum and the teachers' report on the extent to which the professional development was designed to support the policy environment. In general, where correlations existed between familiarity with policy instruments and professional development, correlations also were found between teachers' report of the curriculum reflecting the policy instruments and the professional development. For example, there was a 0.28 correlation between the extent to which professional development helped a teacher meet the state assessment requirements and the extent to which the teacher's curriculum reflected state or local student assessments (Table 59), and a 0.26 correlation between the extent to which professional development helped a teacher meet the state assessment requirements and the teacher's familiarity with the state or local student assessments (Table 58). Either professional development activities that addressed policy instruments increased the likelihood that teachers would integrate the policy instrument into the curriculum, or teachers who integrated the policy instruments into the curriculum were more likely to recognize them as a focus of their professional development.

Table 58. Correlation between familiarity with policy instruments and the extent to which the professional development was designed to support the policy environment or the extent to which the professional development enhanced the teacher's knowledge and skills

	Familiarity							
	Reading				**Mathematics**			
Professional Development	**Cont Stand**	**Curric Frame**	**Student Assess**	**Perform Stand**	**Cont Stand**	**Curric Frame**	**Student Assess**	**Perform Stand**
Well matched to your school's or department's plan to change practice	.21	.16	.17	.16	.18	.16	.14	.15
Designed to support reform efforts under way in your school	.23	.20	.18	.19	.20	.20	.15	.19
Designed to support state or district standards or curriculum frameworks	.27	.19	.24	.22	.27	.21	.22	.22
Designed to support state or district assessment	.26	.19	.22	.21	.26	.21	.20	.20
Helped me adapt my teaching to meet state assessment requirements	.28	.27	.26	.26	.30	.27	.29	.30
Helped me adapt my teaching to meet state standards or curriculum framework requirements	.29	.29	.27	.26	.30	.29	.29	.30
Learned how to help students engage in collaborative inquiry	.20	.23	.19	.20	.21	.26	.23	.24
Gained confidence in using new pedagogical approaches in teaching math	.17	.16	.15	.18	.22	.20	.18	.22
Gained confidence in using new pedagogical approaches in teaching reading/English/ language arts	.22	.24	.21	.26	.22	.23	.18	.25
Feel more motivated to draw from a wide variety of methods when teaching	.22	.26	.21	.25	.24	.24	.22	.26

Table 59. Correlation between the extent to which the policy instrument is reflected in the curriculum and the extent to which the professional development was designed to support the policy environment or the extent to which the professional development enhanced the teacher's knowledge and skills

	Reflected in Curriculum								
	Reading				Mathematics				
Professional Development	**Cont Stand**	**Curric Frame**	**Student Assess**	**Perform Stand**	**Cont Stand**	**Curric Frame**	**Student Assess**	**Perform Stand**	**NCTM Stand**
Well matched to your school's or department's plan to change practice	.22	.23	.20	.19	.20	.19	.21	.18	.15
Designed to support reform efforts under way in your school	.25	.26	.22	.22	.22	.24	.20	.21	.15
Designed to support state or district standards or curriculum frameworks	.27	.27	.25	.24	.25	.25	.16	.22	.13
Designed to support state or district assessment	.24	.26	.25	.22	.25	.25	.15	.21	.15
Helped me adapt my teaching to meet state assessment requirements	.26	.29	.28	.27	.28	.28	.27	.31	.18
Helped me adapt my teaching to meet state standards or curriculum framework requirements	.27	.31	.26	.27	.28	.30	.27	.31	.16
Learned how to help students engage in collaborative inquiry	.17	.22	.18	.21	.20	.23	.20	.26	.16
Gained confidence in using new pedagogical approaches in teaching math	.13	.18	.13	.19	.18	.21	.18	.25	.21
Gained confidence in using new pedagogical approaches in teaching reading/English/ language arts	.18	.24	.20	.24	.16	.18	.17	.25	.14
Feel more motivated to draw from a wide variety of methods when teaching	.21	.26	.22	.25	.23	.25	.21	.27	.14

The integration of content standards, curriculum frameworks, and student assessments into the reading curriculum and the integration of content standards and curriculum frameworks into the math curriculum were correlated with professional development activities designed to support state or district standards or curriculum frameworks and professional development activities designed to support state or district assessments (Table 59).

The integration of the NCTM standards in the curriculum was the least correlated with professional development, while the integration of math performance standards in the curriculum was the most correlated with professional development. We expect that many districts have content and performance standards modeled after the NCTM standards, however. Teachers may be familiar with the state's or district's content and performance standards but unaware of the influence of the NCTM standards in their curriculum. Additionally, elementary teachers may be the least likely of all teachers (elementary, middle, and high) to be familiar with the NCTM standards.

No correlations were found between the *amount* of professional development and the extent to which the policy instruments were integrated into the curriculum.

We also examined the relationship between the reported professional development and the change from 1997 to 1998 in teacher response to the questions about familiarity of policy instruments and the extent to which the policy instruments are reflected in a teacher's curriculum. No correlations were identified for reading or math. For both subjects, the highest correlation was 0.13, with most correlations less than or equal to 0.10.

5.5 Relationship Between Professional Development and Curriculum and Instruction

As described previously, teachers were asked a variety of questions about curriculum and instruction in reading. For the many possible relationships between curriculum and instruction and professional development, only a few correlations were above 0.25. The extent to which teachers use higher achieving students to work with lower achieving students (0.26) and the extent to which teachers emphasize content area reading strategies (0.26 for the lowest achieving students and 0.25 for typical students) were modestly correlated with professional development focused on teachers helping students engage in collaborative inquiry. In general, most of the correlations between curriculum and instruction in reading and professional development were less than 0.15, indicating little relationship between professional development and teacher activities, student activities, or skills emphasized.

Teachers were also asked a variety of questions related to selected mathematical competencies (i.e., memorize facts, understand concepts, solve equations, collect/interpret data, solve word problems, and solve novel problems), their teaching strategies, and student activities. The mathematical competency variable was derived by summing teachers' reports on how much they emphasized a competency over a series of 10 math topics (e.g., using number lines and rulers, operations with fractions, etc.). Teachers were asked to address the math competency item for both typical and low-achieving students.

Unlike reading, there were modest correlations (i.e., $\geq$ 0.25) between professional development and the mathematical competencies. For example, for both typical students and low-achieving students, professional development that helped the teacher learn how to help students engage in collaborative inquiry and gain confidence in using new pedagogical approaches in teaching math were modestly correlated with the extent the teacher emphasized collecting and interpreting data (typical students: 0.30 and 0.29; lowest achieving students: 0.26 and 0.25), and solving novel problems (typical students: 0.31 and 0.31; lowest achieving students: 0.29 and 0.28). Professional development that enhanced teachers' knowledge or skills to help students engage in collaborative inquiry also was modestly correlated with teachers' reported level of preparation to use cooperative learning groups in mathematics (0.28) and with teachers' reported level of preparation to integrate mathematics with other subject areas (0.29).

For typical students, there were also correlations between the extent to which professional development helped the teacher feel more motivated to draw from a variety of teaching methods and the extent to which the teacher's lessons focused on helping students learn to collect and interpret data (0.25) and solve novel problems (0.25) (Table 60). However, most of the correlations between professional development and student activities and teaching strategies were less than 0.15.

Table 60. Correlations between mathematical competencies and the extent to which the professional development enhanced the teachers' knowledge and skills

Professional Development	**Typical Students**						**Lowest Achieving Students**					
	A	**B**	**C**	**D**	**E**	**F**	**A**	**B**	**C**	**D**	**E**	**F**
Helped me adapt my teaching to meet state assessment requirements	.19	.13	.16	.24	.23	.23	.18	.12	.15	.22	.20	.22
Helped me adapt my teaching to meet state standards or curriculum framework requirements	.17	.12	.15	.25	.22	.21	.16	.12	.14	.23	.18	.20
Learned how to help students engage in collaborative inquiry	.24	.10	.22	.30	.27	.31	.22	.08	.20	.26	.23	.29
Gained confidence in using new pedagogical approaches in teaching math	.22	.11	.20	.29	.26	.31	.22	.09	.19	.25	.22	.28
Gained confidence in using new pedagogical approaches in teaching reading/English/ language arts	.19	.11	.16	.24	.24	.26	.18	.11	.16	.21	.20	.23
Feel more motivated to draw from a wide variety of methods when teaching	.17	.15	.16	.25	.22	.25	.14	.13	.15	.22	.19	.23

Key:
A = Memorize facts
B = Understand concepts
C = Solve equations
D = Collect/interpret data
E = Solve word problems
F = Solve novel problems

5.6 Conclusions

Using school districts' policy documents to identify the provisions in place in each LESCP district with regard to standards, assessment, and accountability, we arrayed the 18 districts from high to low along a rough continuum of standards-based reform. This classification is probably valid at the extremes, although it is not precise enough to support fine distinctions. And, indeed, the responses of teachers to questions about standards-based reform did differ in the expected ways across district policy environments: those in high-reform districts were significantly more likely to report familiarity with and adherence to various policy instruments, such as standards, assessments, and frameworks. The changes over time were interesting as well: the greatest amount of change in teacher responses was found in those districts that did not start out at either the high or the low extreme of the policy environments.

Professional development varied a great deal across districts, although only some of the variation was associated with the gross distinction between high- and low-reform environments. One example of a difference was the greater emphasis on learning about assessment in high-reform districts. Another was the greater focus on state and district reforms, whereas the school's own reform plan was more often the focus of professional development in low-reform environments. The overall amount of professional development diminished across the 2 years of the study. Teachers gave mixed reviews to their professional development, with under 50 percent rating the quality as "high" and under 25 percent saying it had helped them in a variety of ways "to a great extent."

Some aspects of professional development were modestly associated with differences in teachers' responses about the policy instruments and about classroom practices. In particular, the focus of professional development showed a few relationships with the skills teachers emphasized in their mathematics curriculum. However, professional development was not discernibly associated with *changes* in practice for individual teachers over the 2 years of this study.

APPENDIX A

This appendix provides further detail on the study's analytic methods. In it, we estimate variances due to schools, teachers, and students for the LESCP longitudinal sample, then describe the hypothesis-testing procedure used to determine whether particular teacher practices were significantly associated with student gains.

The data structure of students, nested within fourth-grade teachers, nested within schools suggests a hierarchical model for the data. There is an enormous amount of literature devoted to models of this form in educational research and in many other fields (Bryk & Raudenbush, 1992). For this analysis, we used a three-level model with student (*i*) nested within class (*j*) nested within school (*k*). We assume the test score difference (fourth-grade – third-grade score), d_{ijk}, for student *i* in classroom *j*, in school *k* follows the model:

$$d_{ijk} = \beta + v_k + u_{jk} + r_{ijk} \tag{1}$$

where β is a fixed effect and the other three components of equation (1) are independent with the following distributions: $v_k \sim N(0, \sigma_v^2)$, $u_{jk} \sim N(0, \sigma_u^2)$, and $r_{ijk} \sim N(0, \sigma_r^2)$, where $N(0, \sigma^2)$ denotes the normal distribution with mean 0 and variance σ^2. Equation (1) specifies a three-level analysis of variance (ANOVA) model of the following type (i.e., Scheffe, 1959):

- Nested (or hierarchical) model: the nested structure of the data described above dictates this assumption.
- Random effects (or components of variance): The three components are assumed to be random, as opposed to fixed. This assumption is that the students, classrooms, and schools are representative of a larger population.

The three random components can be explained as follows:

- School effect (v_k): Adds the same amount to the gain for each student in the k^{th} school that could reflect equipment, leadership, and Title I reforms made at the school.
- Teacher effect (u_{jk}): Adds the same amount to the gain for each student of the j^{th} teacher in the k^{th} school that reflects factors such as the teacher's knowledge, education, and experience.

- Student effect (r_{ijk}): Reflects factors such as student drive, home situation, etc.

It follows from the assumptions that the differences have the normal distribution with mean β and variance given by the sum of the variances of the three random effects. In symbols we have $d_{ijk} \sim N(\beta, \sigma_T^2)$ where $\sigma_T^2 = \sigma_u^2 + \sigma_v^2 + \sigma_r^2$. In this model, we assume the same expected gain for each student independent of classroom, school, and other demographic factors. Thus, the estimate of β will coincide with the mean difference of the longitudinal sample. A histogram of the distribution of differences for the closed-ended math test indicated that the normal distribution provides a good fit in this case; similar results are obtained for the other seven tests and subtests.

We estimate the parameters of equation (1) using standard methodology. The estimates of the variance components are shown in Table A-1, while Table A-2 shows the percentage of variance due to schools, teachers, and students. For example, the percentage of the variance due to the teacher on the closed-ended reading is given by 100*79.2/(19.7+79.2+653.6) = 10.5 percent. Table A-2 shows that the vast majority of the variance is due to the student. The table also shows that the teacher percentage of the variance is slightly lower for reading than for math. The percentage of variance explained by the teacher is comparable to that obtained by Hanushek, Kain, and Rivkin (1998), who found that variation in teacher quality accounts for at least 7.5 percent of student achievement, using scores from a large sample of fourth- through sixth-graders in Texas.

Introducing additional fixed-effect (regression) parameters could reduce each of the variances. In the analytic approach described below, we introduce teaching practices to reduce the variance due to the teacher.

Table A-1. Variance of LESCP longitudinal sample score gains by school, teacher, and student

	Variance		
Test or Subtest	**School**	**Teacher**	**Student**
Reading Closed Ended	19.7	79.2	653.6
Vocabulary	50.1	114.8	1280.9
Comprehension	41.2	111.9	1048.0
Reading Open Ended	185.4	220.4	2641.4
Math Closed Ended	97.5	135.9	638.0
Problem Solving	107.3	93.0	851.9
Procedures	138.0	286.6	1525.9
Math Open Ended	44.2	138.3	774.2

Table A-2. Percentage of variance of LESCP longitudinal sample score gains

	Percentage of Variance		
Test or Subtest	**School**	**Teacher**	**Student**
Reading Closed Ended	2.6	10.5	86.9
Vocabulary	3.5	7.9	88.6
Comprehension	3.4	9.3	87.3
Reading Open Ended	6.1	7.2	86.7
Math Closed Ended	11.2	15.6	73.2
Problem Solving	10.2	8.8	81.0
Procedures	7.1	14.7	78.2
Math Open Ended	4.6	14.5	80.9

The variances shown in Table A-1 are based on the entire LESCP longitudinal sample. However, as explained above, the results for students in classes where the teacher did not complete a questionnaire were not used in estimation of the impact of teaching practices (since we chose not to impute teacher responses). As a check, we also computed the variances for the subset of students in classes where the teacher completed a questionnaire. The results for this subset of students were similar to those shown in Tables A-1 and A-2 (the maximum difference in the student percentages was 3 percent), thus, the results are not shown here.

We describe next the analytic procedure used to investigate the relationship between teaching practices and the change in student test scores.

Following Hill and Goldstein (1998), we used a four-level model with repeated test scores (*t*) nested within student (*i*) within class (*j*) within school (*k*). We assume the test score Y_{tijk} for student *i* in classroom *j*, in school k at time *t* follows the model:

$$Y_{tijk} = \alpha + \beta t + \gamma t T_{jk} + \varepsilon_{tijk} \qquad (2)$$

where α, β, and γ are fixed-effect parameters, ε_{tijk} is a random error with mean value zero, and T_{jk} denotes the teaching practice in classroom *j* in school *k*. With this model, the expected score for student *i* in classroom *j* in school *k* at baseline (*t=0*) is $E(Y_{oijk}) = \alpha$, and in followup (*t=1*) is $E(Y_{1ijk}) = \alpha + \beta + \gamma T_{jk}$, so the expected difference is $E(d_{ijk}) = \beta + \gamma T_{jk}$ where $d_{ijk} = Y_{1ijk} - Y_{0ijk}$. If the teaching practice has no impact on scores, $\gamma = 0$ so that the expected gain is β as in the model described in equation (1) above. However, if the teaching practice is non-zero, students in different classes have different expected gain depending on their classroom teacher.

The teaching technique was treated as a quantitative variable even though many of the responses were of the ordered categorical type with four to six categories. A typical math question is of the form "When you teach operations with fractions, how much do you emphasize memorizing facts?"[7] The four possible responses to this question were the following: no emphasis, occasional emphasis, emphasized moderately, and emphasized a lot. We coded these four responses as 1 to 4 and used these values in equation (2). The model assumes that a change of teaching practices between two adjacent categories (i.e., change from "occasional emphasis" to "emphasized moderately") changes the expected gain of each student in the class by a constant amount.

Because of the hierarchical structure, the random error, ε_{tijk}, of equation (2) can be decomposed in the sum of four terms: a random school component (v_k), a random classroom component (u_{jk}), a random student component (r_{ijk}), and a measurement error component (e_{tijk}) as follows:

$$\varepsilon_{tijk} = v_k + u_{jk} + r_{ijk} + e_{tijk} \tag{3}$$

This decomposition reduces the measurement error variance thereby creating more powerful hypothesis tests. We assume the four components of equation (3) are independent with the same following distributions: $v_k \sim N(0,\sigma_v^2)$, $u_{jk} \sim N(0,\sigma_u^2)$, $r_{ijk} \sim N(0,\sigma_r^2)$, and $e_{tijk} \sim N(0,\sigma_e^2)$. We estimate the three fixed-effect parameters (α,β,γ) and the four variance components ($\sigma_v^2, \sigma_u^2, \sigma_r^2, \sigma_e^2$) as described below. The three random components v_k, u_{jk}, and r_{ijk} have the same interpretation as in the description of equation (1).

Now, we contrast the model described by equations (2) and (3) with the ANOVA model described in equation (1). Whereas equation (1) contains four parameters, the model of this section has seven. The four common parameters are those of equation (1) namely, (σ_v^2, σ_u^2, σ_r^2, β) while the model of this section has three additional parameters (α,γ,σ_e^2). Also, the model here uses both Y_{1ijk} and Y_{0ijk} (rather than only their difference, d_{ijk}). Thus, twice as many dependent variables are used in the estimation of the model parameters of this section. This allows the three additional parameters to be estimated. Whereas equation (1) can be estimated using ANOVA techniques, equation (2) and (3) specifies a mixed model, which is more difficult to estimate.

---

[7] Different math topics are substituted for "operations with fractions" and different competencies are substituted for "memorizing facts."

For LESCP data analysis of the mixed model described in equations (2) and (3), we investigated the following two software packages: Hierarchical Linear Models (HLM, Bryk, Raudenbush, Seltzer, & Congdon, 1988), and PROC MIXED in SAS (SAS Institute, 1996). We decided to use PROC MIXED to estimate the model parameters and to test for significance due to the following advantages:

- Less onerous data processing (special data processing for each analysis stage is not required)
- Allows all the data to be used (in HLM a portion is often discarded)
- No limitations on the number of levels (our version of HLM was limited to three levels)

APPENDIX B

This Appendix contains information on the relationship between student gains in test scores and teacher responses to survey items for students who scored in the bottom quartile nationally on the pretest of the Stanford 9. We analyzed the eight different tests and subtests (four math and four verbal) to determine which ones have a statistically significant relationship (at the .05 level) with questionnaire responses. For each test, we used the pretest results to split the students into a bottom quarter nationally group using the students' pretest results (third-grade scores from 1997).

In the body of this report, we discussed results for students split into the bottom quarter and top three-quarters of their class. Because a disproportionate number of students (40 percent on the closed-ended reading portion of the Stanford 9) in the classes we studied are in the bottom quarter of test results when compared to the nation, we will now report results for students who scored in the bottom quarter nationally. These results illuminate the impact of a variety of teaching practices on student learning for students who, although they may do well in comparison to their classmates, would be identified as low achievers when compared to all students across the nation.

For each of the questions, we estimated the parameters of the model defined in equation (1) and (2) in Appendix A of this report. We tested the hypothesis $\gamma = 0$ and summarized the results in the following tables for both reading scores and math scores. *If the hypothesis was not rejected* at the .05 significance level the *cell is blank* in the table. If the *test was significant* at the .05 significance level, a *plus sign* is shown in the table if more of the quantity is *related to a significant increase* in test scores. Significant *negative relationships* are shown as a *minus* sign.

Table B-1. Curriculum emphasis in reading/language arts *for low-achieving students*: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 9)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Extent of Emphasis, in Teaching Low-Achieving Students, on:				
Comprehension	+			
Vocabulary				
Oral Reading				
Content Area Reading Strategies			-	
Phonics/Word Attack				

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's self-reported emphasis on comprehension in teaching low-achieving students and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended reading test.

Table B-2. Total exposure to student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 12)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Total Exposure (Frequency x Time Per Lesson) for Students:				
Read Materials of at Least One Paragraph	+	+		
Read Aloud	+			
Read Books They Choose Themselves				+
Practice Word Attack				
Read Content Area Materials				
Practice Phonics				
Complete Workbooks/ Skill Sheets		+		+
Talk in Small Groups About What They Have Read	+	+	+	
Write About What They Have Read			-	
Work at a Computer	+	+		+

Table reads: There was a significant (p<.05) positive relationship between (1) total exposure, in a fourth-grade teacher's classroom, to reading materials of at least one paragraph and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended reading test.

Table B-3. Frequency of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 13)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Frequency of Activity for Students:				
Read Materials of at Least One Paragraph	+	+	+	
Read Aloud	+			
Read Books They Choose Themselves		+	+	
Practice Word Attack				
Read Content Area Materials				
Practice Phonics				
Complete Workbooks/ Skill Sheets		+		
Talk in Small Groups About What They Have Read	+			
Write About What They Have Read		-	-	-
Work at a Computer		+		

Table reads: There was a significant (p<.05) positive relationship between (1) frequency, in a fourth-grade teacher's classroom, of reading materials of at least one paragraph and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended reading test.

Table B-4. Duration of student instructional activities in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 14)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Time Spent per Lesson in Activity for Students:				
Read Materials of at Least One Paragraph				+
Read Aloud				
Read Books They Choose Themselves			-	
Practice Word Attack				
Read Content Area Materials			-	
Practice Phonics		+		
Complete Workbooks/ Skill Sheets	-			
Talk in Small Groups About What They Have Read		+		
Write About What They Have Read				
Work at a Computer				

Table reads: There was a significant ($p<.05$) positive relationship between (1) time per lesson, in a fourth-grade teacher's classroom, in reading materials of at least one paragraph and (2) the gains made by that teacher's bottom-quarter nationally students on the vocabulary subtest of the closed-ended reading test.

Table B-5. Instructional strategies in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 17)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Extent of Use, in Teaching Students of Different Achievement Levels, of:				
Extra Time With Low Performers				+
Different Instructional Materials				
Frequent Assessments				+
Heterogeneous Grouping				
Homogeneous Grouping		-	-	
One-on-One Instruction				

Table reads: There was a significant ($p<.05$) negative relationship between (1) a fourth-grade teacher's self-reported use of homogeneous grouping and (2) the gains made by that teacher's bottom-quarter nationally students on the closed-ended reading test.

Table B-6. Teacher preparation in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 19)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
How Well Prepared To:				
Use Small Group Instruction	+		+	+
Take Existing Skills Into Account	+			
Integrate Reading/ Language Arts With Content Areas	+			
Teach Heterogeneous Groups	+		+	
Use a Variety of Assessment Strategies	+			

Table reads: There was a significant (p<.05) positive relationship between (1) a fourth-grade teacher's self-reported level of preparation to use small-group instruction and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended reading test.

Table B-7. Policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 21)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Teacher's Familiarity With:				
Student Assessments	+			
Performance Standards	+			
Content Standards	+			
Curriculum Frameworks				
Extent Reflected in Curriculum:				
Student Assessments	+			
Performance Standards	+			
Content Standards	+			
Curriculum Frameworks				

Table reads: There was a significant (p<.05) relationship between (1) a fourth-grade teacher's self-reported familiarity with content standards, and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended reading test.

Table B-8. Perceived appropriateness of policy instruments in reading/language arts: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 23)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
How Appropriate for the Teacher's Students Are:				
Content Standards				
Curriculum Frameworks				
Student Assessments				
Performance Standards				
Integration with Content Areas				

Table reads: There were no significant ($p<.05$) relationships between (1) a fourth-grade teacher's rating of the appropriateness of policy instruments for his or her students and (2) the gains made by that teacher's bottom-quarter nationally students on the Stanford 9 reading tests.

Table B-9. Parent involvement and students ready to learn (reading): relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 24)

Teacher Report	**Reading Test or Subtest**			
	Open-ended Reading	**Closed-ended Reading**		
		Overall	**Comprehension**	**Vocabulary**
Teacher's Report on:				
For how many of your low-achieving students are parents or guardians at least moderately involved in school activities?		+		+
How many of your low-achieving students usually come to school prepared to learn?	+	+	+	+

Table reads: There was a significant (p<.05) positive relationship between (1) a fourth-grade teacher's response on the number of his or her low-achieving students that come to school prepared to learn and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended reading test.

Table B-10. Topical coverage in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 26)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Number of Lessons Taught in:				
Word Problems With Addition, Subtraction				
Multi-Digit Multiplication	+			
Rounding		+	+	+
Using Number Lines and Rulers		+	+	
Operations With Fractions	+		+	
Finding Length, Perimeter With Pictures			+	
Solving Equations With One Unknown		+	+	+
Distance Problems	+		+	
Determining Central Tendency	+			-
Solving Equations With Two Unknowns			+	

Table reads: There was a significant ($p<.05$) positive relationship between (1) the number of lessons a fourth-grade teacher reported teaching on using number lines and rulers and (2) the gains made by that teacher's bottom-quarter nationally students on the closed-ended mathematics test.

Table B-11. Cognitive demand in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 27)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Average Emphasis (Across Topics) on Teaching Low-Achieving Students to:				
Understand Concepts	+			
Solve Equations	+			
Solve Word Problems			+	
Collect/ Interpret Data			+	
Memorize Facts				
Solve Novel Problems			+	

Table reads: There was a significant ($p<.05$) positive relationship between (1) the average emphasis reported by a fourth-grade teacher on teaching low-achieving students to understand concepts and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-12. Teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 29)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Total Exposure (Frequency x Time Per Lesson) to Teacher Activity:				
Work an Exercise at the Board	-			
Lead Whole-Group Discussions	-	-		
Lecture or Present				
Discuss Multiple Approaches To Solving a Problem		+		
Use Manipulatives				-
Administer a Test		+	+	+

Table reads: There was a significant ($p<.05$) negative relationship between (1) total exposure reported by a fourth-grade teacher to the teacher leading whole group discussions and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-13. Frequency of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 30)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Frequency of Teacher Activity:				
Work an Exercise at the Board		+	+	+
Lead Whole-Group Discussions				
Lecture or Present				
Discuss Multiple Approaches To Solving a Problem			+	
Use Manipulatives				-
Administer a Test				+

Table reads: There was a significant ($p<.05$) positive relationship between (1) frequency with which a fourth-grade teacher reported working an exercise at the board and (2) the gains made by that teacher's bottom-quarter nationally students on the closed-ended mathematics test.

Table B-14. Duration of teacher instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 31)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Time per Lesson Spent in Teacher Activity:				
Work an Exercise at the Board	-			
Lead Whole-Group Discussions		-		
Lecture or Present		-		-
Discuss Multiple Approaches To Solving a Problem	-			
Use Manipulatives				
Administer a Test		+	+	

Table reads: There was a significant ($p<.05$) negative relationship between (1) time per lesson reported by a fourth-grade teacher in lecturing or presenting and (2) the gains made by that teacher's bottom-quarter nationally students on the closed-ended mathematics test.

Table B-15. Student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 34)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Total Exposure (Frequency x Time Per Lesson) to Student Activity:				
Respond Orally to Questions	-			+
Work Individually on Worksheets	-			+
Work in Small Groups	-			
Discuss Solutions in Whole Group	-			
Drill on Computational Skills				
Participate in Student-Led Whole-Group Discussions	-			+
Analysis With Tables and Graphs				
Use Calculators To Solve Problems	-	+		
Assignments Requiring More Than a Paragraph				
Work With Manipulatives				
Assignments Taking More Than a Week	-	+	+	+
Review Completed Homework in Class				+

Table reads: There was a significant ($p<.05$) negative relationship between (1) total exposure reported by a fourth-grade teacher to oral response to questions and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-16. Frequency of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 35)

Teacher Report	Mathematics Test or Subtest			
	Open-ended Mathematics	Closed-ended Mathematics		
		Overall	Problem Solving	Procedures
Frequency of Student Activity:				
Respond Orally to Questions				
Work Individually on Worksheets				+
Review Completed Homework in Class	+	+	+	+
Drill on Computational Skills	+	+	+	+
Work With Manipulatives			+	
Work in Small Groups				
Discuss Solutions in Whole Group		-		-
Assignments Requiring More Than a Paragraph	+		+	
Participate in Student-Led Whole-Group Discussions	-		+	
Analysis With Tables and Graphs				
Use Calculators To Solve Problems				
Assignments Taking More Than a Week		+	+	+

Table reads: There was a significant ($p<.05$) positive relationship between (1) frequency with which a fourth-grade teacher reported that students worked individually on worksheets and (2) the gains made by that teacher's bottom-quarter nationally students on the procedures subtest of the closed-ended mathematics test.

Table B-17. Duration of student instructional activities in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 36)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Time Spent per Lesson in Student Activity:				
Respond Orally to Questions	-			
Work Individually on Worksheets	-			
Review Completed Homework in Class	-			
Drill on Computational Skills	-			
Work With Manipulatives	-			
Work in Small Groups	-			
Discuss Solutions in Whole Group	-			
Assignments Requiring More Than a Paragraph	-			
Participate in Student-Led Whole-Group Discussions				
Analysis With Tables and Graphs				
Use Calculators To Solve Problems	-		+	
Assignments Taking More Than a Week	-			

Table reads: There was a significant ($p<.05$) negative relationship between (1) the time per lesson reported by a fourth-grade teacher for students' oral response to questions and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-18. Teacher preparation in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 39)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
How Well Prepared To:				
Present Mathematics Concepts	+	+	+	
Teach Heterogeneous Groups	+	+	+	+
Manage a Class Using Manipulatives				
Use Cooperative Learning Groups	+			
Use the Textbook as a Resource	+	+	+	
Take Students' Existing Concepts into Account	+	+	+	+
Use a Variety of Assessment Strategies		+	+	+
Integrate Math With Other Subject Areas		+	+	

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's self-reported level of preparation to present mathematics concepts and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-19. Policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 41)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Teacher's Familiarity With:				
Student Assessments	+			+
Performance Standards	+			
Content Standards				
Curriculum Frameworks				
Extent Reflected in Curriculum:				
Student Assessments		+	+	+
Performance Standards	+	+	+	+
Content Standards		+	+	+
Curriculum Frameworks	+			
NCTM Standards	+		+	

Table reads: There was a significant ($p<.05$) positive relationship between (1) a fourth-grade teacher's self-reported familiarity with state or district student assessments and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-20. Perceived appropriateness of policy instruments in mathematics: relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 43)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
How Appropriate for the Teacher's Students Are:				
Student Assessments				
Performance Standards				
Content Standards				
Curriculum Frameworks				
Integration With Content Areas	+	+		

Table reads: There was a significant (p<.05) positive relationship between (1) a fourth-grade teacher's rating of the appropriateness for his or her students of integrating mathematics with content areas and (2) the gains made by that teacher's bottom-quarter nationally students on the open-ended mathematics test.

Table B-21. Parent involvement and students ready to learn (mathematics): relationship between fourth-grade teacher's response and fourth-grade students' gain for *bottom-quarter nationally* students (Compare with Table 44)

Teacher Report	**Mathematics Test or Subtest**			
	Open-ended Mathematics	**Closed-ended Mathematics**		
		Overall	**Problem Solving**	**Procedures**
Teacher's Report on:				
For how many of your low-achieving students are parents or guardians at least moderately involved in school activities?				
How many of your low-achieving students usually come to school prepared to learn?				

Table reads: There were no significant ($p \leq .05$) relationships between (1) a fourth-grade teacher's response on parent involvement for his or her low-achieving students or how many of their students come to school prepared to learn and (2) any gains made by that teacher's bottom quarter students on any of the subtests of the mathematics test.

REFERENCES

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Bryk, A.S., Raudenbush, S.W., Seltzer, M., & Congdon, R. (1988). *An introduction to HLM: Computer program and user's guide* (2^{nd} edition). Chicago: University of Chicago Department of Education.

Diggle, P.J., Liang, K.Y., & Zeger, S.L. (1994). *Analysis of longitudinal data*. Oxford: Clarendon Press.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (1998 Project). *Teachers, schools, and academic achievement*. NBER Working Paper 6691.

Hill, P.W., & Goldstein, H. (1998). Multilevel modeling of educational data with cross-classification and missing identification for units. *The Journal of Educational and Behavioral Statistics, 23*, 117-128.

SAS Institute (1996). SAS/STAT software: Changes and enhancements through Release 6.11, Cary, North Carolina.

Scheffe, H. (1959). *The analysis of variance*. New York: John Wiley.

Stanford Achievement Test Series (Ninth Edition.) *Spring norms book (1997)*. San Antonio: Hartcourt Brace & Company.

Stanford Achievement Test Series (Ninth Edition.) *Technical tables (1996)*. San Antonio: Hartcourt Brace & Company.

Stanford Achievement Test Series (Ninth Edition.) *Urban spring norms (1996)*. San Antonio: Hartcourt Brace & Company.